Ph. D. Dissertation in Economics

Modelling and Measuring Production of Korean Universities

- Knowledge Production Function and Scientific Collaboration -

August 2014

Graduate School of Seoul National University

Technology Management, Economics, and Policy Program

JongWuk Ahn

Modelling and Measuring Production of Korean Universities

- Knowledge Production Function and Scientific Collaboration -

지도교수 이 정 동

이 논문을 경제학 박사 학위논문으로 제출함

2014 년 8 월

서울대학교 대학원

협동과정 기술경영경제정책 전공

안 종 욱

안종욱의 박사 학위논문을 인준함

2014 년 8 월

위 원 장 ______________________________ (인)

부위원장 ______________________________ (인)

위 원 ______________________________ (인)

위 원 ______________________________ (인)

위 원 ______________________________ (인)

부모님께

Abstract

Modelling and Measuring Production of Korean Universities

- Knowledge Production Function and Scientific Collaboration -

Ahn, JongWuk
Technology Management, Economics, and Policy Program
The Graduate School
Seoul National University

This study explores the nature of the knowledge production process. For this purpose, we exploit the scientometric perspective, econometric methods, social network standpoint, and production function approach derived from the endogenous growth theory. All of them combine to

help us better understand the nature of the knowledge production process, with a focus on scientific collaboration.

First, we demonstrated that multi-university collaboration produces higher-impact articles when it includes a top-tier university. It is also found that elite universities experience impact degradation of their scientific results when they collaborate with lower-tier institutions, whereas their lower-tier partners gain impact benefits from the same collaboration. We also revealed that Korean universities are unlikely to work with other universities from the same tier.

Second, we performed social network analysis (SNA), which is appropriate for examining interactive relationships such as co-authorships, to depict structural relations within the scientific collaboration networks of Korean universities. We defined the "external world" as the coupling of domestic non-university institutions with organizations from foreign countries, separating them from the "university world." Then, we have drawn a bigger picture to investigate structural relationships in scientific co-authorship networks for Korean academia.

Third, we developed a model of knowledge production function, embedding a better understanding of the collaborative characteristics in the scientific production process. The analysis was undertaken from a "university world" perspective, based on the hypothesis that universities should be affected by both the "inside" and the "outside" of the university world. It provided an assessment of the scientific returns to investment in university research, particularly considering the existence and characteristics of two spillover effects: One is "cross-university," and the other is "from the US." The results implicate that spillover effect has a real existence and a deep relation with scientific collaboration, while this effect is weaker than the influence of research and development (R&D) expenditures on scientific results. The small spillover effect requires either much greater collaboration or improving the efficiency of co-operation.

Keywords: Knowledge production function · Scientific collaboration · University research · Partner selection · Co-authorship network · Knowledge spillovers

Student Number: 2005-21293

Contents

Chapter 1. Introduction 1

1.1 Science Economics.. 1

1.2 Universities as Engines of Knowledge Economy 1

1.3 Science as an Output 2

1.4 Outline.. 3

Chapter 2. Conceptual Framework 4

2.1 Introduction 4

2.2 Returns to R&D... 5

2.3 Science and Technology Studies 8

Chapter 3. Scientific Impact and Partner Selection 13

3.1 Introduction 13

3.2 Data and Methodology 20

3.3 Empirical Results 28

3.4 Conclusion.. 42

Chapter 4. Social Structure of Scientific Network........................... 44

4.1 Introduction 44

4.2 Data and Methodology 47

4.3 Network Development and Divided Network......................... 52

4.4 Structure of Scientific Network.. 58

4.5 Relating Node Characteristics to Scientific Production 67

4.6 Conclusion: Partner selection strategy for more scientific production 73

Chapter 5. Knowledge Production Function for Korean University 75

5.1 Introduction 75

5.2 Data 78

5.3 Methodology 88

5.4 Empirical Results 99

5.5 Conclusion.. 113

Chapter 6. Conclusion.. 116

6.1 Summary of Results 116

6.2 Implications.. 118

6.3 Contribution and Limitation.. 120

Bibliography.. 122

Appendix 1: Deriving the Basic Knowledge Production Function.... 136

Abstract (Korean)... 138

List of Tables

Table 1 Increasing collaboration by subject fields (WCs) 30
Table 2 Marginal advantage of alliances between universities by school tier (2001–2009) 36
Table 3 Fraction of papers in tier combinations and propensity ratio for tier pairings 38
Table 4 Data sources 48
Table 5 Descriptive data of the scientific collaboration network 53
Table 6 Types of university in terms of their partner selection 66
Table 7 Centrality scores for academic network 69
Table 8 Descriptive statistics of centralities, 2006–2010 70
Table 9 Correlation among network centralities and scientific production in the world of universities 71
Table 10 Descriptive statistics of partner preference for Korean universities, 2006–2010 72
Table 11 Correlation among scientific production and partner preference 72
Table 12 Research expenditure at Korean universities, 2008–2012 81
Table 13 Variables for knowledge production function 85
Table 14 Descriptive statistics: research outputs, knowledge stock, and control variables 87

Table 15 Results from basic knowledge production function (dependent variable log ‘sci’) 100
Table 16 Results from basic knowledge production function (dependent variable log ‘arti’) 101
Table 17 Knowledge production function including share of funding source (dependent variable log ‘sci’) 103
Table 18 Knowledge production function including share of funding source (dependent variable log ‘arti’) 104
Table 19 Knowledge production function including network centralities (dependent variable log ‘sci’) 106
Table 20 Knowledge production function including core preferences (dependent variable log ‘sci’) 107
Table 21 Correlation among variables 107
Table 22 Extended knowledge production function for Korean universities (dependent variable log ‘sci’) 109
Table 23 Elasticities of returns to own and others’ R&D 111

List of Figures

Figure 1 Trend in share of authorship structures 28
Figure 2 Trend in fraction of multi-university cooperation by number of authors 29
Figure 3 Percentage of the number of universities participating in collaborative research 31
Figure 4 Probability that a scientific article belongs to the high-impact papers (2001–2009) 33
Figure 5 Marginal advantage of between-school collaboration compared to within-school teamwork (2001–2009)............ 35
Figure 6 Marginal impact advantages for cross-tier scientific alliances 41
Figure 7 Scientific collaboration network development at Korean universities.. 52
Figure 8 Scientific collaboration network, 2010 54
Figure 9 Collaboration network for Korean universities, 1981–2010. 55
Figure 10 Scientific network during 2006–2010................................ 58
Figure 11 The scientific network constructed from selected edges (> 300) for 2006–2010. 59
Figure 12 D.NU and some of its dependents (sky blue nodes) 60
Figure 13 Universities focusing on D.NU and USA (sky blue nodes) 61
Figure 14 After removing D.NU and USA (> 300)............................ 62

Figure 15 Universities independent from the external world (> 300). 63
Figure 16 Degree centrality of the scientific collaboration network ... 67
Figure 17 Betweenness centrality of the scientific collaboration network 68
Figure 18 Closeness centrality of the scientific collaboration network 69
Figure 19 Scattered plot of residuals from estimating the basic knowledge production function (dependent variable log 'sci') 98

Chapter 1. Introduction

1.1 Science Economics

There are several reasons why economists should pay attention to science. The first is that science is a major source of economic growth. While the lags from research to economic growth may be long, science's impact on the economy is incontestable, and the evidence is quite a real phenomenon. The second reason is that scientific results have the properties of public good. Economists have concerns regarding the failure of economies (or markets) to produce public goods. The third reason is that the public property of science and its spillovers are fundamental to the economic concept of endogenous growth theory (Romer, 1990). In conclusion, "the foremost reason economists have for studying science is the link between science and economic growth." (Stephan, 2010).

1.2 Universities as Engines of Knowledge Economy

Universities are organizations that perform a fundamental role within contemporary societies through education and other methods of disseminating knowledge (Perkmann et al., 2013). Universities are among the oldest institutions, having governed not only to adapt to frequent external shocks, but also to expand its size and roles (Ben-David, 1977). As we move towards a more knowledge-intensive economy, universities face pressure to link their research more closely to economic necessities. However, despite the increased pressures on the third mission, university research will undoubtedly survive and continue to evolve as they have done over previous centuries (Martin, 2012). In the economic literature, many studies have analyzed university–industry relationships, including technology transfer and intellectual property rights (Organisation for Economic Co-operation and Development [OECD], 2002).

1.3 Science as an Output

While science economics has an interest in economic value for basic research, another discipline is concerned about the process of science that combines streams of work focusing on the history of science, philosophy of science, and sociology of science (Fagerberg et al., 2012b).

Some works have used a production function approach to model scientific production, which uses some inputs to yield scientific outputs or knowledge such as the publication of scientific journal articles (Crespi, 2007; Crespi & Geuna, 2008).

In this study, we explore the nature of the knowledge production process. For this purpose, we exploit the scientometric perspective, econometric methods, social network standpoint, and production function approach from endogenous growth theory. All of them complementarily help us better understand the nature of the knowledge production process with a focus on scientific collaboration.

1.4 Outline

This study consists of six chapters, and the remainder is organized as follows. Chapter 2 describes the conceptual framework for this study. It provides an overview of the economic returns to research and development (R&D), the knowledge spillover effect, scientific collaboration, and the social network analysis (SNA). Chapter 3 conducts a descriptive analysis to show the growth of collaborative research at Korean universities over the last three decades. It also analyzes the impact advantages of collaboration by school tier, the marginal advantages of between-school collaboration over within-school teamwork, and partner selection in terms of propensity ratios. Chapter 4 defines the "external world" as the coupling of domestic non-university institutions with organizations from foreign countries, and draws a bigger picture to investigate structural relationships in scientific co-authorship networks for Korean academia. In Chapter 5, a model of knowledge production function embedding a better understanding of collaborative characteristics is developed. Finally, Chapter 6 summarizes the results from this study, provides policy implications, and states the contributions and limitations of this study.

Chapter 2. Conceptual Framework

2.1 Introduction

The problem of this study that of economic modelling and measuring the production of Korean universities, might be solved by a knowledge production function approach from growth accounting (Solow, 1957; Abramowitz, 1986; Romer, 1990).

In this chapter, we introduce a conceptual framework through this thesis from now on. As the goal of this study is the development of a model that is suitable for reflecting actual status, we borrow some perspectives from a new discipline that of science and technology studies (STS) literature. Specifically, constructs named “scientific collaboration” and “social network” are used to obtain certain intuitions for modelling.

2.2 Returns to R&D

"The large contribution of research and development to economic growth is an undisputed fact today." — Foray (2004, p. 51)

Research and development (R&D) can raise productivity by enhancing the quality or lowering the average production costs of current goods (or services) or by simply expanding the options of final goods (or services) or intermediate inputs available. Returns to R&D can be defined as returns to investment in research and development and other innovation assets, and are generally an interesting subject for accountants, managers, policy makers, and economists (Hall et al., 2010). Economists have been developing assorted methodologies to estimate the rate of return to R&D expenditure. Most literature has used a familiar growth accounting framework developed with measures of R&D investment or R&D capital, essentially relating the growth of total factor productivity (TFP) to R&D expenditure. To put it another way, the residual factor in production growth that is not explained by ordinary inputs (labor, capital, and intermediate inputs) is considered the result of R&D yielding innovation.

A pioneering study by Griliches (1979) delivered the structure of the economic measurement of returns to R&D in a production function context. In this, two major issues are discussed: The measurement of output when plenty of R&D investment is funded for quality enhancement or nonmarket goods (e.g. scientific articles); and the measurement of the stock of R&D capital.

2.2.1 Knowledge spillovers and relational capital

The empirical literature on the knowledge production function concentrates on R&D spillovers and extends the Griliches knowledge production framework to measure productivity gains (or other intermediate outputs such as scientific achievement) that arise from knowledge spillovers of R&D investment (Eberhardt et al., 2013). R&D carried out in one unit may yield spillover effects in other units. Such spillovers are made all the more possible and significant if the interacting units are closely related (Hall et al., 2010).

Knowledge is a rival but only partially excludable good in economics. Units will get higher absorptive capacities and more knowledge spillover will take place when more knowledge is codified. This concept of knowledge spillover is crucial for an innovation system that expects growth and development. Note that a knowledge spillover is different from technology transfer. While technology transfer refers to the trade in technology, a non-pecuniary spillover refers to an unintended knowledge transfer, in which payment is not involved (Hall et al., 2010).

If one starts from the concept of pure physical space, the precondition for knowledge spillovers is the physical proximity to firms in the same sector, to firms in different sectors, and to universities and research institutes, the typical places where knowledge can be produced. Physical proximity increases the probability of contact between economic actors, thereby allowing knowledge to spread more easily and produce useful spillovers. On the other hand, if one takes into account the concept of relational space, the precondition for the creation of knowledge spillovers becomes the cultural proximity of actors. This cultural proximity is the basis for the existence of relational capital (Capello & Faggian, 2005). Relational capital is therefore the “substratum” of collective learning, precisely as physical space is the necessary condition for “traditional” knowledge spillovers.

2.3 Science and Technology Studies

According to Fagerberg et al. (2012b), “the knowledge society” has three strands of study as a knowledge base: (a) Innovation studies (Fagerberg et al., 2012a); (b) entrepreneurship studies (Landström et al., 2012); and (c) science and technology studies (Martin et al., 2012).

Among the others, science and technology studies (STS) is one of a number of new research fields to emerge over the last four or five decades. It has interests on the history, philosophy, and sociology of science (Martin et al., 2012). Among the variety research fields within STS, we have concentrated on scientific collaboration and the social network perspective to frame the knowledge production process of Korean university schools.

2.3.1 Scientific collaboration

The exploration of knowledge production has been conducted by scientists from various research areas over several decades. Of particular interest is scientific collaboration, which has become a mainstay of knowledge creation (Katz & Martin, 1997; Sonnenwald, 2007). Previous studies have illustrated the relationship between the increasing dominance of scientific collaboration and the scientific impact of new knowledge (Beaver, 2004; Guerrero-Bote et al., 2013; Wray, 2002).

Many scholars have examined the increase in collaborative research. Cronin et al. (2003) illustrated the growing importance of collaboration in psychology and philosophy literature during the 20th century. Adams et al. (2005) found that the number of authors on a scientific paper written at American universities increased by 50% between 1981 and 1999. Gazni et al. (2012) discovered that multi-entity publications increased worldwide during 2000–2009. Additionally, not only is scientific collaboration on the increase, it has been shown to benefit collaborators and the research itself. The observed advantages of collaborative research include mentoring or teaching (Collins, 1974), varied insights from different disciplines (Hoch, 1987), better research productivity (Lee & Bozeman, 2005), and the improved quality of research results (Franceschet & Costantini, 2010).[1]

Aside from these merits, citation impact is the most popular and frequent measure of the benefit of cooperative research. For example, Beaver (2004) quantitatively affirmed that collaborative research has greater epistemic authority (which correlates with citation impact) than research performed individually. Many studies, such as Guerrero-Bote et al. (2013), have used scientific impact to analyze the benefits of international scientific collaboration. Using a functional explanation for the persistence of scientific collaboration, Wray (2002) argued that cooperation helps scientists

1 Collaborative research is also regarded as useful one from a technological standpoint. For example, Lee (2008) posited that the quality and value of patents are related with research collaboration.

successfully access the resources they require, enabling a more effective realization of epistemic research goals.

2.3.2 Social network analysis

As the interaction among entities becomes an important factor, innovation studies require social network analysis (SNA), which investigates the relationship between entities (Smart et al., 2007). In particular, the structure and position of nodes and the evolution mechanisms of social networks could prove interesting. The two main streams of literature on SNA are from sociology and physics. By abstracting the relationships of actors to a collection of links between nodes, SNA offers a graphical methodology for understanding the structural characteristics of social phenomena. This approach is helpful in such areas as when designing efficient communication networks (Monsuur, 2007) or effective peer-to-peer social networks (Wang & Sun, 2008).

Sociologists are concerned about the position of agents in networks and about the network structure itself. To measure these, they developed network indices and coefficients such as centralities, cliques, structural equivalence, and distance (Brass, 1984; Wasserman & Faust, 1994). The social structure of a network has a major implication on the information flow and the diffusion of innovation. The notion that social cohesion (cooperation among colleagues) and structural equivalence (competition among rivals) influences the diffusion of medical innovation has been argued for long time (Burt, 1987). According to Everard and Henry (2002), companies bridging e-commerce companies with established leading firms are linked to high performance. Korfiatis et al. (2006) measured the degree centralities of keywords and authors in Wikipedia to investigate the contributions of authors and the heterogeneity of co-editions.

While sociologists are interested in the roles of units and their relationships within network structures, physicists have examined the internal structure and network's evolution mechanisms (Kim, 2008). From this perspective, social phenomena in gross society are explained as being formed by the simple motivation of each member. For the last decade, it is shown that many real networks are statistically heterogeneous, being different from the Erdos and Reiny model. Barabási and Albert (1999) found that the degree distribution of nodes follows a power law, and referred to this sort of network as "scale-free," while Watts and Strogatz (1998) found some networks having few nodes and a huge number of links in a short path relating any two nodes, and named these "small world networks."

Establishing a network of collaboration that utilizes knowledge is a principle of the innovation process (Freeman, 1991). SNA provides techniques that examine innovation activities related with the cooperation of members and the exchange of knowledge and resources. Some indicators that investigate the innovation network are centrality coefficients, centralization, clustering coefficients, network density, and structural equivalence (Scott, 1991; Wasserman & Faust, 1994). They could consider both the individual properties and environment, and both the organization's structure and its performance (DeBresson & Amesse, 1991).

Social networks consist of actors (or nodes, vertices, etc.), ties (or links, edges, etc.), and the forms of them that can be varied according to the purpose of a study. The actor may be either an individual or a collective unit. Firms, governments, and agencies are examples of collective unit actors, and this thesis focuses on one of them, the universities within Korean academia. Actors make ties that link to others within social communities. The definition of ties here is in specific relations such as agreements, contracts, communications, and collaborations; the forms of ties can be either directed or non-directed, and can be either binary or valued. Directed ties are created when an actor is sending and their partner is receiving, and non-directed ties are created when the actors are not designated as either senders or receivers. Valued ties indicate multiple counting between a pair of actors, while binary ties do not have counts for the relationship (Wasserman & Faust, 1994). Normally, actors of equivalent type and ties are investigated; however, sometimes one or more type of actors and ties are presented in a network, as this study shows. The selection of types of actors and ties in an analysis is dependent on theoretical concerns and research objectives.

Chapter 3. Scientific Impact and Partner Selection[2]

3.1 Introduction

Exploration of knowledge production has been conducted by scientists from various research areas for several decades. Of particular interest is scientific collaboration, which has become a mainstay of knowledge creation (Katz & Martin, 1997; Sonnenwald, 2007). Previous studies have illustrated a relationship between the increasing dominance of scientific collaboration and the scientific impact of new knowledge (Beaver, 2004; Guerrero-Bote et al., 2013; Wray, 2002).

Many scholars have examined the increase in collaborative research. Cronin et al. (2003) illustrated the growing importance of collaboration in psychology and philosophy literature during the 20th century. Adams et al. (2005) found that the number of authors on a scientific paper written at American universities increased by 50% between 1981 and 1999. Gazni et al. (2012) also discovered that multi-entity publications increased worldwide during the period 2000–2009. Additionally, not only is scientific collaboration

---

[2] This chapter is based on the published paper authored by JongWuk Ahn, Dong-hyun Oh, and Jeong-Dong Lee (Ahn et al., 2014). The title is "The scientific impact and partner selection in collaborative research at Korean universities" from the journal *Scientometrics*.

on the rise, it also has been shown to benefit collaborators and the research itself. Observed advantages of collaborative research include mentoring or teaching (Collins, 1974), varied insights from different disciplines (Hoch, 1987), better research productivity (Lee & Bozeman, 2005), and improved quality of research results (Franceschet & Costantini, 2010).[3]

Aside from these merits, citation impact is the most popular and frequent measure of the benefit of cooperative research. For example, Beaver (2004) quantitatively affirmed that collaborative research has greater epistemic authority (which correlates with citation impact) than research performed individually. Many studies, such as Guerrero-Bote et al. (2013), have used scientific impact to analyze the benefits of international scientific collaborations. Using a functional explanation for the persistence of scientific collaboration, Wray (2002) argued that cooperation helps scientists successfully access required resources, enabling the more effective realization of epistemic research goals.

Meanwhile, for the past several decades, science and technology management researchers have focused on research and development (R&D) alliance issues. Partnership subjects have been particularly popular in knowledge management literature. In many R&D alliance studies, counterparts in scientific alliances have been regarded as a key factor for

[3] Collaborative research is also regarded as useful one from a technological standpoint. For example, Lee (2008) posited that the quality and value of patents are related with research collaboration.

improving a firm's performance. The partner effect, or nature of the relationship, has been shown to influence R&D alliance outcomes (Saxton, 1997). The type of R&D partner (e.g., competitors, suppliers, customers, or universities and research institutes) also impacts a firm's performance (Belderbos et al., 2004). Research strands on alliances regard partner selection as a particularly important strategic activity. Van der Valk et al. (2010) studied inter-organizational collaboration among Dutch life science firms in order to better understand the partner selection process for inter-organizational R&D collaboration. Later, Diestre and Rajagopalan (2012) developed a theoretical framework to explain the decision making process when selecting partners.

While scientometrics studies have examined the benefits of collaborative research, collaborative issues such as partner selection have received little attention in scientific knowledge production literature. Although some studies have examined collaboration types or patterns, they did not investigate preferences or patterns at partner selection for the collaborative research. Sooryamoorthy (2009) reported that citation counts received by a research paper vary not only depending on whether the research is collaborative, but also by type of collaboration (e.g. international, external-institutional, internal-institutional, or domestic). Gazni and Didegah (2011) compared the influence of collaboration on the citation impact of publications by categorizing collaboration patterns into intra-institutional, inter-institutional, domestic, and international collaboration, and found that the number of institutions involved in a publication positively correlated to the number of citations received. These studies on scientific collaboration have limits that they merely concern the benefits as results of cooperative research rather than examining partnership tendency at the foundation stage of scientific collaboration. Consequently, strategic decision making in scientific cooperation such as partner selection has received little attention. To our knowledge, only Jones et al. (2008) has investigated a tendency to choose a scientific alliance in multi-university collaborations, where a tendency toward social stratification was observed.[4]

This study attempts to bridge the gap between scientometrics (specifically scientific collaboration), and knowledge management (specifically partner selection), by examining scientists' preferences when choosing research collaborators, and linking this process to the goal for the scientific collaboration, i.e., marginal citation impact. To accomplish this, we analyzed a large data set of scientific research articles written by authors from 213 Korean universities.

We first conducted a descriptive analysis to show the growth of collaborative research at Korean universities over the last three decades, using bibliometric data covering all subject fields in science, technology, and engineering between 1981 and 2010, provided by the Web of Science's (WoS)

---

[4] In their research, Jones et al. (2008) found the social stratification, a tendency that universities are likely to cooperate within their own tier, in multi-university research collaboration at the USA.; and this stratification has increased over time.

Science Citation Index Expanded (SCIE) feature. To our knowledge, our study covers the longest time span and the largest range of disciplines at Korean universities, and thus addresses the limitations of previous studies examining knowledge production in Korea.

To perform an in-depth investigation of partner selection at Korean universities, we disaggregated the schools into four tiers based on epistemic authority, as measured by total number of citations received. Since our analyses are restricted to similar institutions in one small country using a single language, we have likely eliminated all potential considerations for partner selection other than scientific specialty.[5] The scientific ability to produce successful research, therefore, can be regarded as the key factor driving collaborator selection among Korean university scientists. Consequently, our data selection validates tier disaggregation, and confirms that the issue of scientific partner selection at the institutional level can be adequately examined by school tier.

After looking into the rise in scientific collaboration from various angles, we analyzed the impact advantages of collaboration by school tier. The probability that a scientific article belongs to the high-impact papers was

[5] When searching for possible partners for scientific alliances, considerations may be scientific specialty (Frenken 2002), language difference (Narin et al., 1991), geometric distance (Hoekman et al., 2010), or national characteristics (Gazni et al., 2012; Leydesdorff & Sun, 2009). However, we can exclude these but the scientific ability by focusing to only universities in a single country.

calculated and compared across authorship structures. Moreover, the marginal advantages of between-school collaboration over within-school teamwork were estimated across tiers using an econometric technique. Finally, partner selection was analyzed in terms of propensity ratios, as expressed by ratios of the observed frequency and the hypothetical random rates for tier combinations. After doing so, we observed a ‘cross-tier’ preference phenomenon; that is, Korean university scientists were unlikely to select scientific partners from same-tier universities. The motive underlying this phenomenon will, of course, be clarified.

The first contribution of this chapter is adding wide-range, long-term description of knowledge production for Korean academia, specifically concentrating on scientific collaboration among universities. In terms of authorship structure, multi-university collaboration has increased while intra-university teamwork and single-author research has declined. Secondly, this chpater supplements empirical evidence of the beneficial effects of collaboration. Alliances between universities yield stronger impact than intramural teamwork if partnerships include elite universities. Top-tier institutions, however, show a negative marginal advantage in citation impact when engaging in collaborative research, whereas schools from other tiers exhibit positive benefits from between-school collaboration. Lastly, the main contribution of this chapter is to examine the foundation phase of scientific collaboration from the standpoint of strategic partner selection. We analyzed

Korean universities' propensity for tier matches in two-university collaborations and found that cross-tier cooperation is preferred to within-tier cooperation. Based on this finding, we suggest that lower-tier universities need to strategically select partners for scientific alliances to enhance the impact of knowledge created, as this appears to be the main cause for inter-tier preference in scientific collaboration between separate schools.

The remainder of this chapter is organized as follows. Section 3.2 explains our data set, key terms, and methodology. Section 3.3 provides results from descriptive and other empirical analyses. The final section presents a brief conclusion.

3.2 Data and Methodology

3.2.1 Data

We used the 1981–2010 South Korea NCR (National Citation Report), a subset of the WoS database for the empirical investigation. The database includes bibliographic and citation information on 297,658 regular scientific articles published between 1981 and 2010 that have at least one author with a Korean address. Our sample focuses on a set of papers written by university scientists and belonging to the SCIE database, including the WoS Subject Categories (WC) of science, technology, and engineering.[6]

In the original data, each paper lists affiliation addresses of all authors. This, however, makes huge difficulties in cleaning information on affiliations since the original authors might have used their own methods of translating Korean addresses to English. To solve this problem, we re-translated the name variations of Korean universities into our own university codes using custom algorithms. By filtering all papers published by 213 Korean universities, we acquired 149,457 articles in 171 WCs between 1981 and 2010.

[6] *The 1981–2010 South Korea NCR* database contains information on the articles from other disciplines like social science, arts, and humanities. However the portion of these fields is very small in the sample for this chapter. We used only science and technology fields, but their share is large enough (97.78%) to be representative.

3.2.2 Authorship structure and team size

Authorship structure, also referred to as the type of collaboration, is a popular criterion for classifying scientific collaboration. Kim (2006) used four different types of authorship structure to examine changes in distribution of scientific papers written by Korean physicists between 1982 and 2000: international collaboration, institution-external collaboration, institutional-internal collaboration, and no collaboration. In order to investigate the frequency of collaborations between scientists at different US universities, Jones et al. (2008) compared the number of sole-authored research papers and the amount of research papers published by collaborators at the same university. As mentioned above, Sooryamoorthy (2009) found that the number of citations received by a research article is affected by not only existence of collaboration but also the types of collaboration. Gazni and Didegah (2011) estimated the impact of the number of authors, institutions, and foreign nations on the number of citations by investigating research papers and books published by Harvard University from 2000 to 2009. They found that the numbers of authors and institutions have a significantly positive correlation with the number of citations, whereas the number of foreign collaborating countries does not.

To describe the knowledge production characteristics or patterns at Korean universities, we categorized authorship structure into three types: solo, within-school collaboration, and between-school collaboration. Team size was defined as the number of authors working on a paper. The solo authorship structure refers to papers are written by a single author. Hence the team size is one. Within-school collaboration indicates scientific research articles produced by multiple authors belonging to the same university. The team size of this structure is larger than one, and the number of collaborating schools equals one. Between-school collaboration signifies research articles co-authored by a number of scholars from various universities. The number of participating schools and the team size are both larger than one. From a scientific alliance perspective, between-school collaboration between two universities is especially useful for examining partner selection preferences through a rate ratio (also known as relative risk in the health sciences).

The variable “team size” indicates the number of authors in an article. Team size has grown over time and is generally regarded as being helpful in gaining higher number of citations. Adams et al. (2005) analyzed 2.4 million scientific articles written by 110 US universities between 1981 and 1999 and found that team size increases the influence of scientific results. Franceschet and Costantini (2010) identified a positive relationship between the number of authors and the number of citations received as well as judgments from peer reviewers using the national research assessment of Italian universities. In the present study, we used regression analysis to control for the effect of team size when estimating the marginal advantage of multi-university collaboration.

3.2.3 Rank and tier of school

To further investigate the characteristics of Korean universities' collaborative research activities, we disaggregated the institutions into four tiers based on epistemic authority, which is measured by the total number of citations received by within-school publications from each university in the corresponding period. Here, "within-school publications" include sole-authored papers and within-school collaborating papers. The consideration of these two authorship structures is sufficient for examining university rankings (Jones et al., 2008). The top 4%, the top 10%, and the top 22% were used to define the boundaries between the four different tiers. These values for % criterion have been selected by iterative experiments with authors' intuition on Korean academia.

3.2.4 Citation impact

The main result of scientific collaboration is the production of new knowledge, and the quality of a scientific article is measured by the number of citations received. According to Beaver (2004), collaborative research possesses more significant epistemic authority than research conducted by an individual, and the epistemic authority is associated with the number of citations gained, probability of citation, and citation lifetime. Guerrero-Bote et al. (2013) used the scientific impact indicator to analyze the benefits derived from international scientific collaboration and concluded that the more countries involved, the higher impact gained. For this chapter, we determined high-impact papers and marginal citation impacts of various authorship structures based on citation counts.

3.2.4.1 High-impact paper

To analyze certain issues on citation impact, we defined articles receiving more than the average number of citations in the same publication year and WC as high-impact papers. An indicator for whether a publication has high impact was used to calculate the probability that a paper would earn above-average citations; this indicator was also used as the dependent variable of the regression analyses in estimating the marginal citation impact advantage.

3.2.4.2 Marginal citation impact advantage

We used regression analysis to estimate the impact advantage of collaborations. The regressions were linear models in which a dummy variable for a high-impact paper is regressed on an indicator variable for authorship structure, subfield (WC), team size, and publication year.

3.2.5 Propensity for tier combination

For between-school collaborations, we focused on collaborative works in which only two universities participated. Two-university cooperation facilitates the examination of the nature of multi-university partnerships. This is especially true for partner selection, which reveals universities' tendency to choose tiers from which their scientific counterparts are.[7] To accomplish this, we first estimated the expected rates of randomly matched collaborations using bootstrapping and then compared those rates with the actual frequency of tier combinations. The ratio of two probabilities, frequently referred to as the rate ratio or relative risk[8], is a popular way to measure the effect of a difference between two outcomes. Those two outcomes or circumstances are actual-matching frequency and expected random-matching rates. Jones et al. (2008) referred to this ratio as the "propensity ratio," and used it to illustrate that multi-university collaborations in the US "are increasingly stratified by in-group university rank."

3.2.5.1 Probability of two-university paper

[7] In the technology management literature, the dyadic perspective is preferred to investigate alliance-structure issues. See Gulati (1998) for the outline of strategic alliances.

[8] For the details about some effect-size measures such as the rate ratio, see Fleiss and Berlin (2009). In the health sciences discipline, the ratio of two probabilities is referred to as the relative risk or risk ratio.

When we defined the probability that multi-university collaborative research includes a university from tier j as P_j, the probability that a two-university paper includes tiers j and k under random matching, P_{jk}, is $P_j \cdot P_k$ if j equals k, and $2 \cdot P_j \cdot P_k$ if j does not equal k.

$$P_{jk} = \begin{cases} P_j \cdot P_k & (j = k) \\ 2 \cdot P_j \cdot P_k & (j \neq k) \end{cases} \tag{3.1}$$

3.2.5.2 Expected frequency

We estimated expected frequency using non-parametric resampling of our two-university sub-sample. The following is our algorithm for a non-parametric bootstrap:

1. Sample n observations randomly without replacement, and obtain a bootstrap data set.
2. Count the number of each tier-matching pair, and calculate the sample rate. For tier j, that rate is P_j^*.
3. Calculate the sample probability that a two-school collaborative work includes tiers j and k, P_{jk}^*, with P_j^* and P_k^* from step 2.

4. Repeat steps 1–3 a large number of times (we chose 1999[9]) and calculate the average of sample rates $P_{jk}^{*,1}$, … , $P_{jk}^{*,1999}$ to obtain the expected rate P_{jk}^{E} for all combinations of j and k.

We obtained, conclusively, 10 P_{jk}^{E} s from P_{11}^{E} to P_{44}^{E}.

3.2.5.3 Propensity ratio

We refer to the ratio of the actual frequency of a given tier combination, P_{jk}^{A}, to its expected frequency in the random-matching situation, P_{jk}^{E}, as the propensity ratio for the given tier paring, R_{jk}.

$$R_{jk} = \frac{P_{jk}^{A}}{P_{jk}^{E}} \tag{3.2}$$

This ratio of two probabilities is referred to the propensity ratio (Jones et al., 2008). If it is greater than unity, an actual tier match is preferred to a counterfactual one under a random-matching scenario; *vice versa*.

[9] Note that repeating for 1999 iterations allows one to easily calculate confidence intervals for common significance levels, e.g. 99% (Carpenter & Bithell, 2000).

3.3 Empirical Results

3.3.1 Increase in teamwork

We explored trends in collaborative research using various descriptive techniques.

Figure 1 illustrates the trend of authorship structure during the study period (1981–2010). Within-school collaboration is a dominant authorship structure throughout the whole period, although its share has decreased from 77% in 1981 to 63% in 2010. Solo papers also decreased steadily, halving from 13% to 6%. Only one authorship structure, between-school collaboration, increased its portion from 10% to 31%. In the earliest years of the study period, from 1981 to 1985 except 1983, solo papers outnumbered between-school collaborations; after 1985, however, solo constitutes the smallest among the three authorship structures.

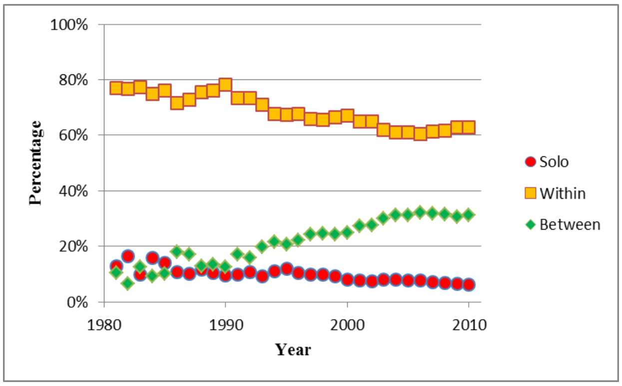


Figure 1 Trend in share of authorship structures

Figure 2 shows that the incidence of publications generally grew regardless of team size. After disaggregating multi-authored papers by the number of authors, we calculated the fraction of multi-university collaborations for each team size. Multi-university collaboration increased not only in the larger-group research outputs (more than three authors) but also in papers written by small groups (two or three authors).

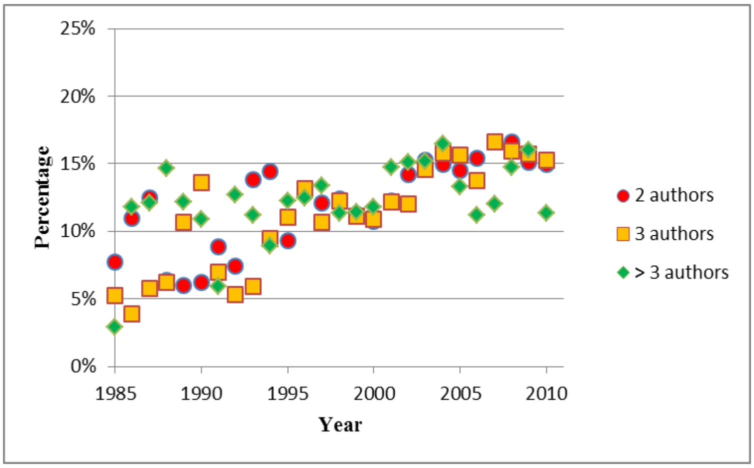


Figure 2 Trend in fraction of multi-university cooperation by number of authors

To examine the rising patterns of collaborative research further, we compared the average fraction of collaboration articles in the first five years studied (1981–1985) with that of the last five years studied (2006–2010) for each subject field in science, technology, and engineering. The rise in within-school teamwork is a phenomenon seen in more than half of the subjects, yielding a 57.9% share (99 out of 171). In addition, 95.9% of the subject fields of science, technology, and engineering demonstrate an increasing share of collaborative research between universities when the first five years of our study are compared with the last five years (164 out of 171).

Table 1 Increasing collaboration by subject fields (WCs)

Total subfields	Subfields in which collaboration has risen			
	Within-school collaboration		Between-school collaboration	
171	99	(57.9 %)	164	(95.9%)

Before ranking the universities and defining the school tiers for further investigation, we examined the number of schools participating in each collaboration. Figure 3 presents the share of the number of universities in collaboration during the study period. With the exception of two-university projects, collaboration across multiple universities increased over time. Although their share has decreased, two-university collaborations represent the most significant portion of scientific research paper production. This dominant portion validates the assumption that two-university collaboration is an acceptable measure of the characteristics of multi-university collaborative scientific research.

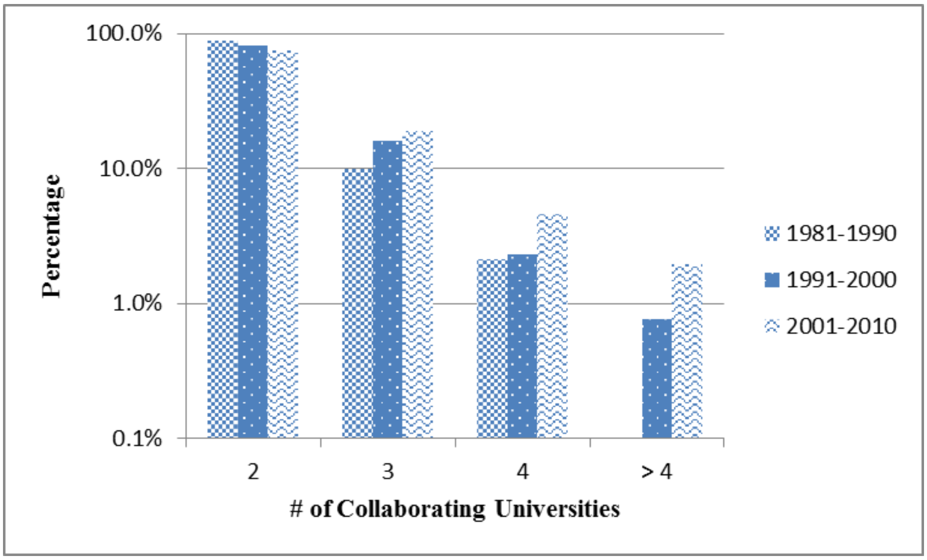


Figure 3 Percentage of the number of universities participating in collaborative research

3.3.2 Impact advantage of collaboration

We calculated the probability that an article gains more citations than the average in the same year and within the same subject field during the period between 2001 and 2009. The citation information in 2010 was not used because it lacks integrity. We compared the probability across authorship structures with respect to school tiers. Regardless of authorship structure, the higher tier had a stronger likelihood of high impact. Collaborative papers within all four tiers also had a higher impact than single-author articles. These results are depicted in Figure 4.

The reason for higher citation rates of collaborative papers has been a matter of interest in literature. Wray (2002) and Beaver (2004) noted the epistemic merits, Katz and Martin (1997) posited that the visibility of a paper causes the more citations earned. These factors may contribute to higher quality of the publication and lead to higher citation counts.

To investigate the impact of between-school collaboration in depth, we categorized collaboration between universities into two additional collaboration types: between with higher tiers and between with lower tiers. Higher tiers refers to tiers I and II, and lower tiers indicates tiers III and IV. The results are depicted in Figure 4. Among between-school collaborations, collaborative research with higher-tier schools shows higher probability than collaboration with lower-tier universities, and this phenomenon spans all tiers. Collaborative work within a tier I school has a probability of 0.37, while collaboration between a tier I university and a lower-tier institution records 0.33. This signifies that a tier I school's within-school collaborative research is more likely to be high-impact than its between-school collaboration with a lower-tier counterpart. In tiers II, III, and IV, the authorship structures of between-school collaboration with both higher and lower tiers exhibit higher probabilities of high impact than other authorship structures, such as solo authorship and within-school collaboration.[10]

[10] The high-impact probabilities for between-school collaboration with tiers III and IV are 0.31, 0.30, and 0.29 at tiers II, III, and IV, respectively; those numbers are greater than the values for within-school collaboration (0.30, 0.28, and 0.25, respectively).

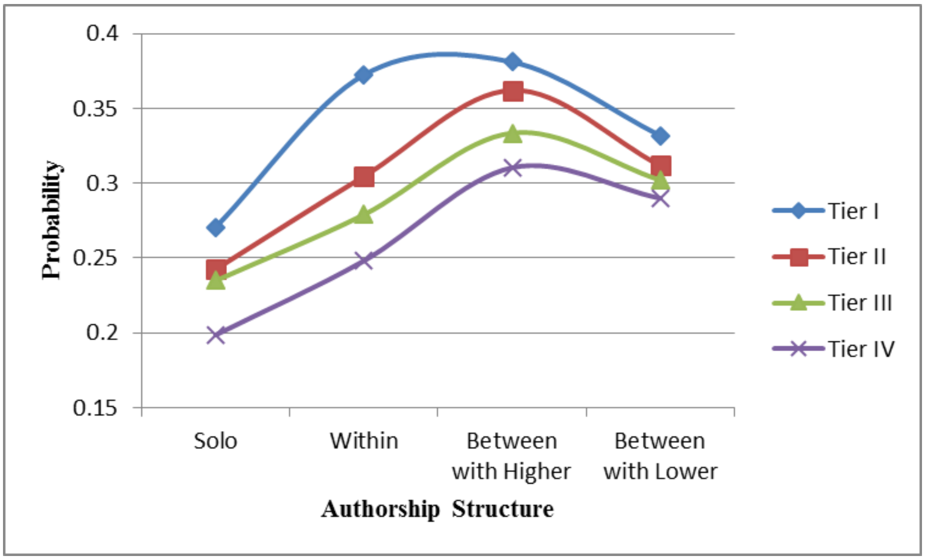


Figure 4 Probability that a scientific article belongs to the high-impact papers (2001–2009)

Note: Solo and Within represents sole-authored papers and papers resulting from collaboration within a school, respectively. Between with Higher (or Lower) indicates between-school collaboration with a partner university from tiers I and II (or tiers III and IV).

To calculate the marginal advantage in citation impact of collaboration between schools over collaboration inside a single university, we ruled out influence of subject field (WC), team size, and publication year using regression models.11 Figure 5 decomposes the marginal advantage of multi-university alliances by tier, in which each bar indicates a separate panel for

[11] Literature has stressed that these factors potentially influence citation counts. Waltman et al. (2011) posited that the effects of research fields and year has to be controlled for calculating indicators of citation impact. The author size is regarded as a positive predictor for highly cited papers (Adams et al., 2005; Franceschet & Costantini, 2010). Therefore, studies based on citation impact are likely to consider all these factors for obtaining specific influences.

each tier. In the 2001–2009 period, collaborative research between two Korean universities was 0.98% less likely to be of high-impact than within-school teamwork.[12] For tier I specifically, the marginal advantage of between-school partnership is –3.41%, meaning that between-school teamwork in tier I is likely to have a lower scientific impact than collaboration within a single university. On the other hand, universities not belonging to tier I have positive values for the marginal advantage of between-school partnerships (2.11%, 3.02%, and 4.42% for tiers II, III, and IV, respectively). Thus, for universities in tiers II, III, and IV, collaboration between schools is more likely to receive more citations than within-school teamwork. Additionally, the marginal advantage increases as school rank decreases; the lower-tier schools gained the more impact from between-school collaborative research.

We further disaggregated the marginal advantage of between-school collaboration over within-school teamwork by the tier of the counterpart institution. In Table 2, each row (hereafter, row-tier) indicates a separate panel for a given tier, and each column (hereafter, column-tier) represents a partner of the row-tier. Numbers in this table provide the degree of marginal advantage over within-school collaborations where universities in the row-

[12] This finding is contrary to that of study by Jones et al. (2008), which is that US universities gain more marginal impact from between-school collaboration than they get from inside-school teamwork. This distinction inspired us to be interested in the reasons behind Korean universities' scientific collaboration.

tiers benefit from collaborations with schools in the column-tiers. In the case of tier I schools, cooperation with schools from other tiers has a statistically significant marginal disadvantage in citation impact over within-school team research. This signifies that tier I schools' collaboration with other tiers leads to a loss in the impact of the scientific knowledge created. Schools in tiers II, III, and IV show positive advantages of collaboration over within-school teamwork for cooperation with a tier I school. Schools in tiers III and IV also show positive marginal advantages of cooperation with another school in tier II. Tier IV universities obtained a positive impact advantage from collaborative research with tier III universities. All the aforementioned marginal advantages and disadvantages are statistically significant at least at a 10% level.

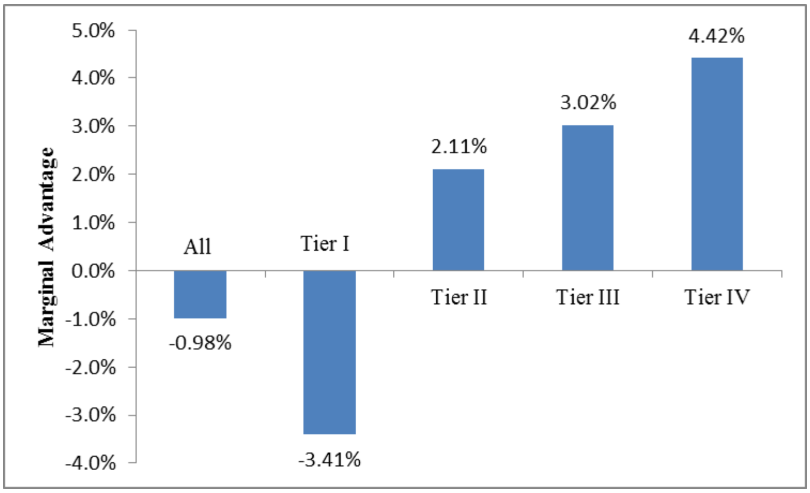


Figure 5 Marginal advantage of between-school collaboration compared to within-school teamwork (2001–2009)

Note: Each bar represents a separate panel for a given tier.

Table 2 Marginal advantage of alliances between universities by school tier (2001–2009)

Tier of School	Tier of Partner			
	Tier I	Tier II	Tier III	Tier IV
Tier I	0.86%	–2.35% ***	–4.92% ***	–5.98% ***
Tier II	4.50% ***	0.80%	0.34%	–3.68% ***
Tier III	4.44% ***	3.55% ***	0.43%	0.49%
Tier IV	7.29% ***	3.10% **	4.33% ***	0.39%

Note: Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ Each value indicates a separate regression result. Schools in the row-tier receive a marginal advantage of between-school collaboration over within-school collaboration when they select a partner from the column tier.

3.3.3 Partner selection in scientific alliance

As mentioned above, we calculated the propensity ratio for tier combinations by dividing the actual frequency of papers by the expected frequency. The expected fractions were obtained using non-parametric resampling. The results are listed in Table 3. The left panels of this table display the actual shares of papers in all tier pairings over three sub-periods. Schools in tier I, which encompasses the top 4% of Korean universities, collaborated on almost 60% of the multi-university publications during the entire research period: 59.6% from 1981–1990, 62.0% from 1991–2000, and 62.3% from 2001–2010. The proportion of tier I schools did not change over time, but the proportion of publications co-authored by a tier IV institution (column sum, actual frequency) diminished rapidly: 62.9% from 1981–1990, 43.3% from 1991–2000, and 30.1% from 2001–2010. This signifies that the-lowest-ranking universities have been alienated from cooperative research between schools.

The right panels of Table 3 exhibit the propensity ratios for tier combinations. As stated previously, a ratio of greater (or less) than unity indicates that the actual fraction of a given tier match is greater (or less) than the corresponding expected value. Intra-tier collaboration was less common than expected in every period and every tier, indicating Korean universities' disinclination to collaborate with schools in the same tier. For example, the actual share was 66% less than expected in tier IV-tier IV partnerships during 1981–1990. However, this tendency appears to have weakened over time, yielding propensity ratios of intra-tier collaboration greater than 0.7 in all tier matches during 2001–2010. This propensity for inter-tier collaboration can be seen as a distinctive feature of Korean universities' knowledge production, and is highly strategically interpreted, especially with regards to alliance selection.

Table 3 Fraction of papers in tier combinations and propensity ratio for tier pairings

	Actual Frequency				Propensity Ratio			
	Tier I	Tier II	Tier III	Tier IV	Tier I	Tier II	Tier III	Tier IV
(1981–1990)								
Tier I	8.7%	7.1%	13.1%	30.7%	0.83	0.66	1.21	1.40
Tier II		1.6%	5.6%	16.9%		0.54	0.98	1.47
Tier III			1.1%	11.3%			0.39	0.99
Tier IV				4.0%				0.34
(1991–2000)								
Tier I	9.9%	12.5%	19.4%	20.2%	0.83	0.92	1.30	1.21
Tier II		3.5%	8.6%	10.6%		0.91	1.01	1.12
Tier III			2.8%	8.1%			0.60	0.77
Tier IV				4.4%				0.76
(2001–2010)								
Tier I	12.8%	18.2%	18.4%	12.9%	1.03	1.09	1.09	1.07
Tier II		4.5%	11.6%	7.5%		0.79	1.03	0.93
Tier III			4.4%	7.6%			0.77	0.93
Tier IV				2.1%				0.72

3.3.4 Motive of cross-tier scientific collaboration

When scientists choose a partner for collaborative research, they consider many factors, such as scientific ability, national differences, institutional characteristics, geographical proximity, and language. Frenken (2002) divided collaboration rationales into economic and intellectual benefits, and described the latter as follows: "Collaboration is intellectually required when specialized knowledge and skills are distributed among different persons." It is obvious that scientific abilities such as research specialty or epistemic significance are major considerations in scientific partner selection. Meanwhile, Gazni et al. (2012) noted that scientifically developed countries are more likely to collaborate with other countries, but added that other elements, such as culture and politics, can also affect collaborative behavior.

At the national level, co-authored publications across various economic sectors have been used to indicate the Triple Helix (university, industry, and government) model for studying knowledge-based economies (Leydesdorff & Sun, 2009). On the other hand, Hoekman et al. (2010) showed that physical distance is a barrier to collaboration, and that territorial borders affect the level of co-publication. Katz (1994) also found that research collaboration decreases exponentially as physical distance increases.

At the same time, linguistic, historical, and cultural factors were found to affect the degree of international co-authorship (Narin et al., 1991). In spite of various other factors affecting partner selection, we were able to focus our analysis on the universities' epistemic power. Since our analyses are restricted to a single institutional sector (university) in one small country using a single language, the only differentiating factor between institutions is scientific specialty. The desire for a successful research result, therefore, can be regarded as the strongest factor inducing Korean university scientists to collaborate. Consequently, Korean universities' preference for cross-tier cooperation over intra-tier teamwork can be explained by considering both the propensity ratio and the marginal citation impact advantage.

Figure 6 depicts the marginal advantage of inter-tier collaborations, which links citation impact to partner selection. Statistically significant figures from Table 2, which lists marginal advantages of inter-tier collaborations, were used for this purpose. In inter-tier university alliances, all lower-tier schools enjoyed significant advantages in citation impact; higher-tier schools, however, seemed not to experience a significant gain.

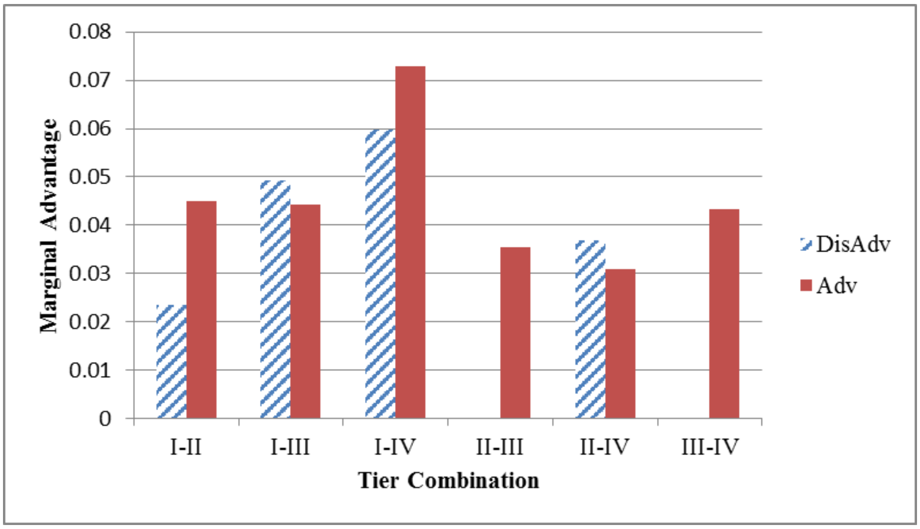


Figure 6 Marginal impact advantages for cross-tier scientific alliances

Note: Adv (or DisAdv) represents the positive (or negative) value of the marginal advantage. Statistically insignificant values are excluded. Left bars represent marginal disadvantages of cooperation if a higher-tier school works with a lower-tier school. Similarly, right bars represent marginal advantages of collaboration if a lower-tier school forms an alliance with a higher-tier school.

Taking both the marginal advantages and the propensity ratios into consideration, lower-tier schools appear to exploit strategic partner selection in order to benefit from scientific alliances. This may suggest that the willingness of lower-tier schools to produce better scientific outputs is stronger than higher-tier schools' reluctance to work with lower-tier counterparts. Especially for universities in Korea, scientific success is not only the key factor for increasing collaborative research but also the main reason for the higher propensity for inter-tier cooperation.

3.4 Conclusion

This chapter examined the characteristics of knowledge production investigating scientific journal articles published by Korean universities from 1981 to 2010. Above all, scientific collaboration between universities is described as the cross-tier cooperative tendency that schools prefer a different level of publication ranking over the same level for partners in collaborative research. Although the reasons for inter-tier partner selection might vary, scientific motive seems to be the key factor driving Korean universities to cooperate with others.

Papers from between-school cooperative research grew in terms of volume during the study period, whereas publications from other authorship structures (solo and within-school collaboration) did not. Regardless of the number of authors (team size), multi-university research outcomes rose. The increase in multi-university collaboration spanned almost every discipline. Although two-school alliances continue to form the majority of collaborations between universities, the number of participating schools also grew.

In terms of citation counts received, collaborative papers performed better than single-author papers. Additionally, cooperation with a higher-tier school yielded a higher citation impact than collaboration within a single university. Tier I universities, ranking in the top 4% of universities by publication citation and being involved in 60% of multi-university articles, sacrificed their citation impact for the advantage in the marginal impact of schools from other tiers (tiers II, III, or IV).

The propensity ratios of tier combinations revealed that Korean scientists prefer inter-tier to intra-tier collaboration. This tendency could be the result of strategic partner selection by lower-tier universities, with the ultimate end goal of achieving scientific success by generating higher-impact research.

As mentioned, our viewpoint is restricted to only universities to determine patterns of and reasons for collaborative research. However, the domain of partner selection for scientific alliances can be extended to domestic industries and governments. In addition, international collaborative research can be studied. Although broadening the scope of this research may be complicated, it will undoubtedly lead to a deeper understanding of scientific collaboration strategies. Another limitation of this chapter is that we did not investigate disadvantages or costs of collaborative research. We believe that additional consideration of benefits and costs in future research will further enhance collective understanding of the scientific knowledge production process.

Chapter 4. Social Structure of Scientific Network

4.1 Introduction

While chapter 3 gave us an interesting picture of scientific collaboration between Korean universities, it lacks some information on the knowledge production process of Korean universities. Collaboration with external to university world is examined seriously in this chapter. For this purpose, we have defined the “external world” as the coupling of domestic non-university institutions with organizations from foreign countries, discriminating them from the “university world.” Then we have drawn a bigger picture to investigate some of the structural relationships in scientific co-authorship networks for Korean academia, consisting of 141 universities and 115 external units.

The generic economic and social benefits from universities have long been regarded as an important source of industrial innovation (Cohen et al., 2002; Mansfield, 1991; Pavitt, 1991; Salter & Martin, 2001; Welsh et al., 2008). At the national level, co-authored publications across various economic sectors have been used to explicate the Triple Helix (university, industry, and government) model for studying knowledge-based economies (Leydesdorff & Sun, 2009).

Selecting a partner from foreign countries for collaborative research has long been a concern of the literature on science and technology studies (STS). Guerrero-Bote et al. (2013) used scientific impact indicators to analyze the benefits derived from international scientific collaboration, and concluded that the more countries are involved, the higher the impact gained. Gazni et al. (2012) noted that scientifically developed countries are more likely to collaborate with other countries.

This chapter aims to supplement the previous chapter with the deeper structure of the collaborative tendencies of Korean university academia. For this purpose, we have exploited the same data used in chapter 3. We took the social network analysis (SNA), which is appropriate for examining interactive relationships such as co-authorships, to depict structural relations. Structural relations are regarded as an independent variable for expecting the innovative performance of actors (e.g. Ahuja, 2000).

The results of this chapter imply that the scientific collaboration network within the Korean university sphere has some characteristics such as: (a) A divided network that indicates less collaboration beyond national borders; (b) A core/peripheral network that reveals the huge importance of a few of the cores; (c) A star network that demonstrates fragmented collaboration among Korean universities. Partner selection for improving scientific output is discussed in the context that networks have those peculiarities.

The remainder of this chapter is organized as follows. Section 2 introduces the data set and the methodology used for analysis and describes it. Section 3 and 4 respectively present the overview and the detailed view of the scientific collaboration network with the Korean university view. The resultant taxonomy of Korean universities is provided in Section 4. Section 5 presents the resultant characteristics of universities, discusses them, and links them to the performance of schools. The last section concludes with the consideration of partner selection strategies for Korean universities.

4.2 Data and Methodology

In this section, we describe the structure and source of our dataset, the variables comprising the knowledge production function, and descriptive statistics for the variables used.

4.2.1 Data sources

The sources of data fall into two groups, depending on whether they are related to personnel and financial information or whether they refer to inter-university collaboration.

The first one, personnel and financial information, is from the website named the "*Higher Education in Korea" service*[13] (hereafter HEiK), provided by the Korean Council for University Education, which allows users to search for information on Korean universities. From this source, we excavated data about the university system by institution for the 2008–2012 period.

The other source of our data is the *1981–2010 South Korea NCR*[14] (hereafter NCR), which was used in the earlier chapter. This source, as a subset of the *Web of Science* (WoS) database, includes the bibliographic and

---

[13] Provisions in 6th section of the *Act on Information Disclosure of Educational Institutions* has obligated the educational institutions to publicly disclose information and to be regulated relevant details. The Web site address for the *Higher Education in Korea service* is www.academyinfo.go.kr.

[14] NCR is an abbreviation for "national citation report."

citation information of 297,658 regular scientific journal articles that have at least one Korean address among the addresses for authors. From this source, we excavated data about collaboration structures for the 1981–2010 period.

Table 4 Data sources

	HEiK	NCR
Information	University related (financial, labor, etc.)	Collaboration related (Bibliometrics)
Source	Korean Council for University Education	Web of Science (Thomson Reuters)

4.2.2 Social network analysis

4.2.2.1 Fundamentals

Social network analysis (SNA) is an interdisciplinary methodology that has mainly been used by sociologists in recent decades. Many social sciences have emphasized the attributes of actors as the key principle in determining their behavior. For example, a theory of utility maximizing principals assumes that an individual makes decisions rationally in accordance with their personal preferences rather than out of concern for others' behavior or their relationship with others. Meanwhile, the network perspective comes from the belief that inter-entity relationships within a network are the most valuable variables for explaining complex social events. SNA has attracted substantial attention from disciplines, because it successfully addresses research challenges that concern both actors' behavior and their performance, which those other traditional social-science perspectives fail to capture. In this situation, a wide range of social issues are investigated with SNA (Wasserman & Faust, 1994).

SNA has three major underlying assumptions originating from distinctive theoretical grounds. First, structural relations among actors are prioritized over the attributes of the actors (gender, age, size, type, etc.) in describing social phenomena. Second, social networks largely influence the beliefs, visions, and behavior of actors. Various structural mechanisms embedded in a society capacitate the realization of social networks. Third, networks are not static, but dynamic as interactive works among actors that change over time (Knoke & Yang, 2008).

SNA has been broadly adopted in science, technology, and innovation (STI) policy research as other social science studies. SNA, more than others, is quite a useful tool for STI policy research, since that interactive relationship is the key of innovative activities based on evolutionary economics and innovation studies. Therefore, a range of STI policies were studied with SNA (Choi, 2010). Many scholars in this strand are concerned about what networks in STI consequences are, and try to explore this question. Structural relations are regarded as independent variables for expecting the innovative performance of actors (e.g. Ahuja, 2000).

4.2.2.2 Concepts

Social networks consist of actors (or nodes, vertex, etc.), ties (or links, edges, etc.), and their forms that can vary according to the purpose of study. Actors may be either individual people or a collective unit. Firms, governments, and agencies are examples of collective unit actors, and this thesis focuses on one of them, the universities within Korean academia. Actors make ties linking to others within their social communities. The definition of ties here is within specific relations such as agreements, contracts, communications, and collaborations; the forms that ties take can be either directed or non-directed, and can be either binary or valued. Directed ties are created when an actor only sends and their partner only receives, and non-directed ties exist when neither actor is designated as sender or receiver. Valued ties indicate multiple counting between a pair of actors, while binary ties do not have counts for relations (Wasserman & Faust, 1994). Normally, the actors of equivalent type and ties would be investigated; however, sometimes one or more types of actor and tie are presented in a network, as in this study. The selection of types of actors and ties in an analysis is dependent on theoretical concerns and research objectives.

4.2.2.3 Measures

SNA offers various methods and measurements to examine the structural relations, properties, and set of regularities in networks. They allow for behavior theory and innovation theory to be understood in descriptive and quantitative terms. In this study, we have focused on analytical concepts such as centrality. One of the major concerns in SNA literature is the search for central and prominent actors. Central actors that have many ties are entities that can depict the whole network, and they can be found intuitively by network figures. The concept of centrality and prestige is used for quantifying the degree of centrality with simple mathematics, and is extensively concerned with innovation studies.

The foundation of the conceptual meaning of centrality is that a central actor, one with the highest value of centrality, is the most active and influential within a network. In general, the higher centrality actors have, the higher their survival rate and the more profit they capture. Some measurements of centrality have been developed: Degree, closeness, and betweenness. Degree accounts for the number of ties, closeness is concerned with how near an actor is to others; and betweenness measures the extent to which an actor mediates between others as a broker in their network (Wasserman & Faust, 1994). This is not only for theoretical concerns, but also for the types of data by which the selection of measurement is made. In addition to these centrality indices for individual actors, a centralization index for the entire network is able to evaluate the extent of network dispersion.

4.3 Network Development and Divided Network

At first, we examine the trends of the dynamics of the whole network of scientific collaboration for Korean universities. We shall investigate the number of total publications of leading-journal articles, publications that have resulted from collaborative research, participating university schools, participating foreign countries, and relationships (co-authorships or links in the network).

4.3.1 Growing network

Figure 7 and Table 5 show the trends in the dynamics of scientific collaboration networks for Korean university spheres. As we can see, all figures have increased.

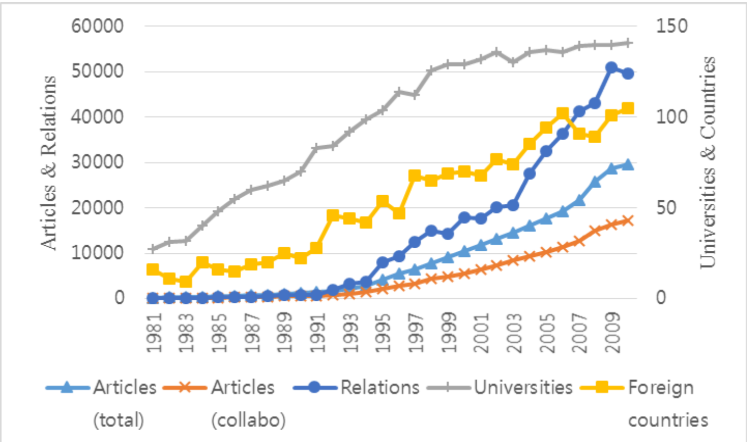


Figure 7 Scientific collaboration network development at Korean universities

Note: The left vertical axis is for the number of articles and relations, the right is for universities and countries.

Table 5 Descriptive data of the scientific collaboration network

	Articles (total)	Articles (collab.)	Universities	Foreign countries	Relations
1981	205	80	27	16	185
1982	246	92	31	11	130
1983	333	133	32	9	189
1984	363	143	40	20	197
1985	485	185	48	16	274
1986	552	229	55	15	337
1987	717	298	60	19	372
1988	878	329	62	20	492
1989	1024	452	65	25	711
1990	1156	477	70	22	780
1991	1363	650	83	28	883
1992	1648	784	84	46	1947
1993	2217	1097	92	44	3282
1994	2765	1441	99	42	3768
1995	4062	2062	104	54	7924
1996	5523	2771	114	47	9289
1997	6380	3371	112	68	12493
1998	7848	4308	126	65	14904
1999	9077	4846	129	69	14327
2000	10446	5470	129	70	17981
2001	11777	6394	132	68	17628
2002	13096	7234	136	77	20121
2003	14521	8383	130	74	20544
2004	16090	9416	136	85	27456
2005	17596	10306	137	94	32465
2006	19186	11345	136	102	36428
2007	21785	12689	139	91	41226
2008	25860	14988	140	89	42990
2009	28804	16330	140	101	50906
2010	29543	17143	141	105	49720

4.3.2 The “university world” and the “external world”

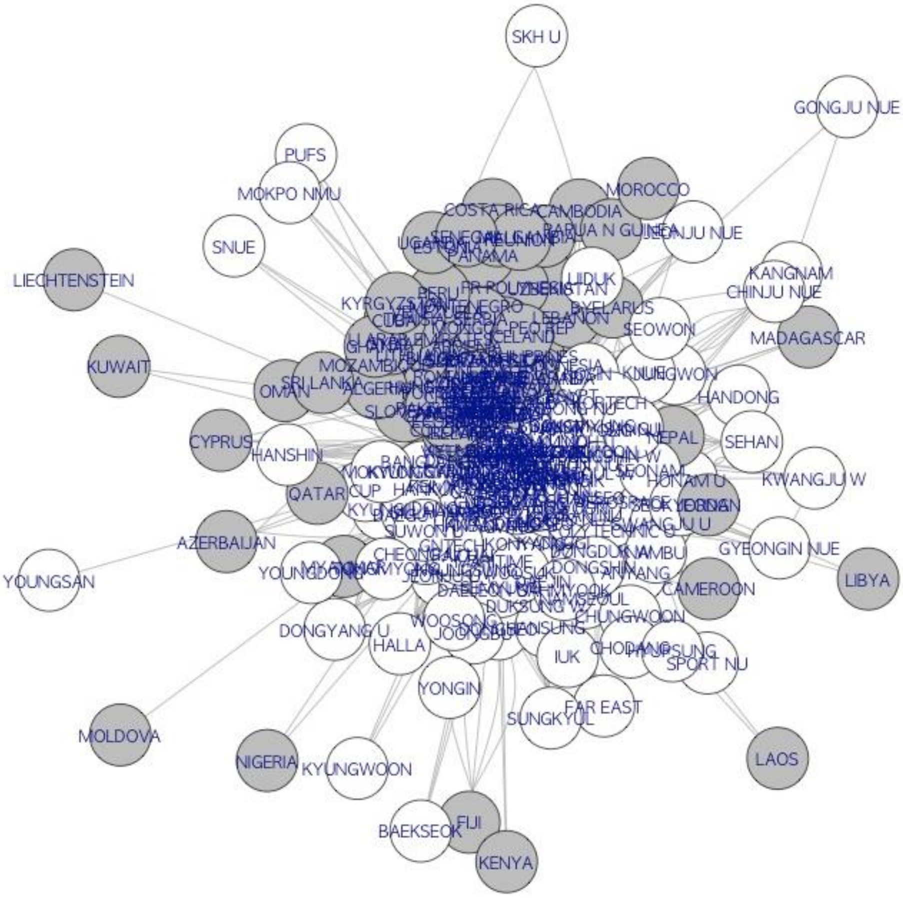
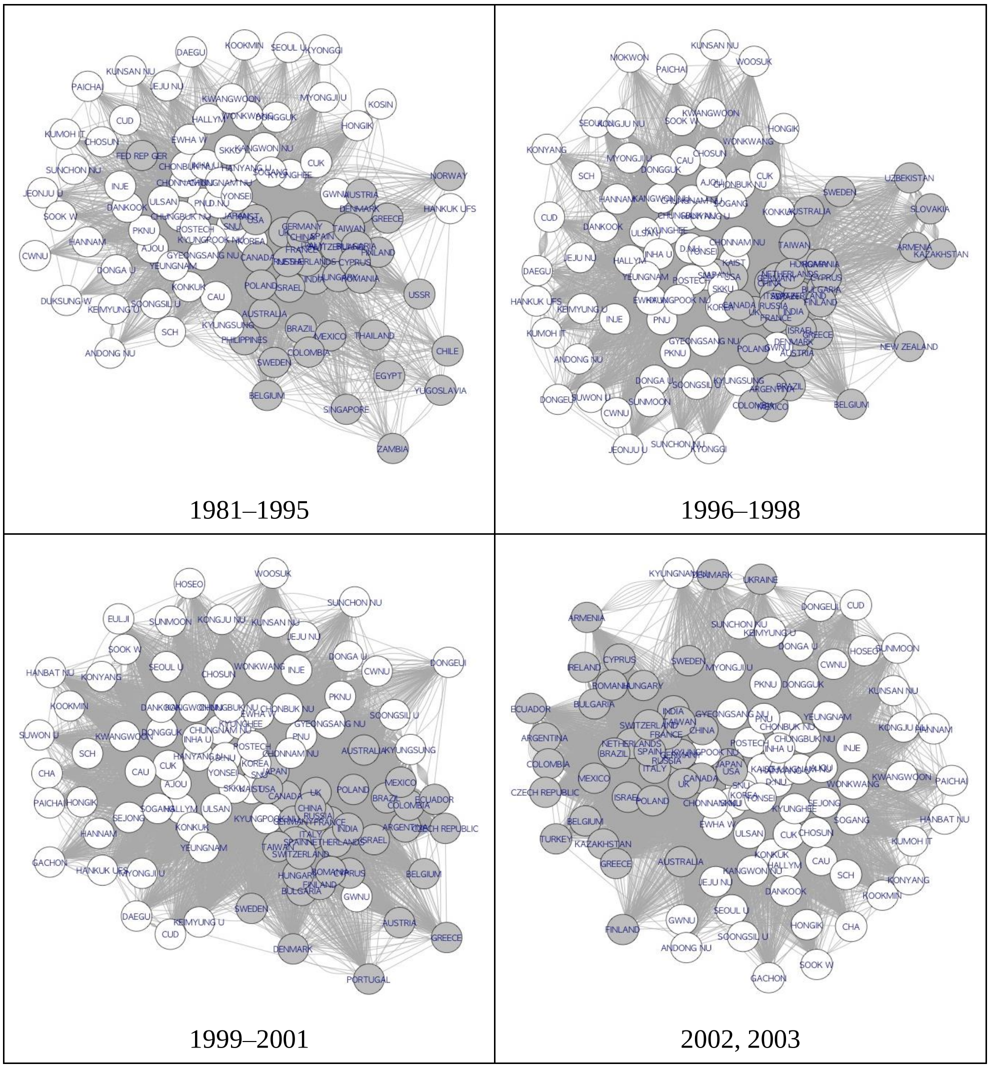
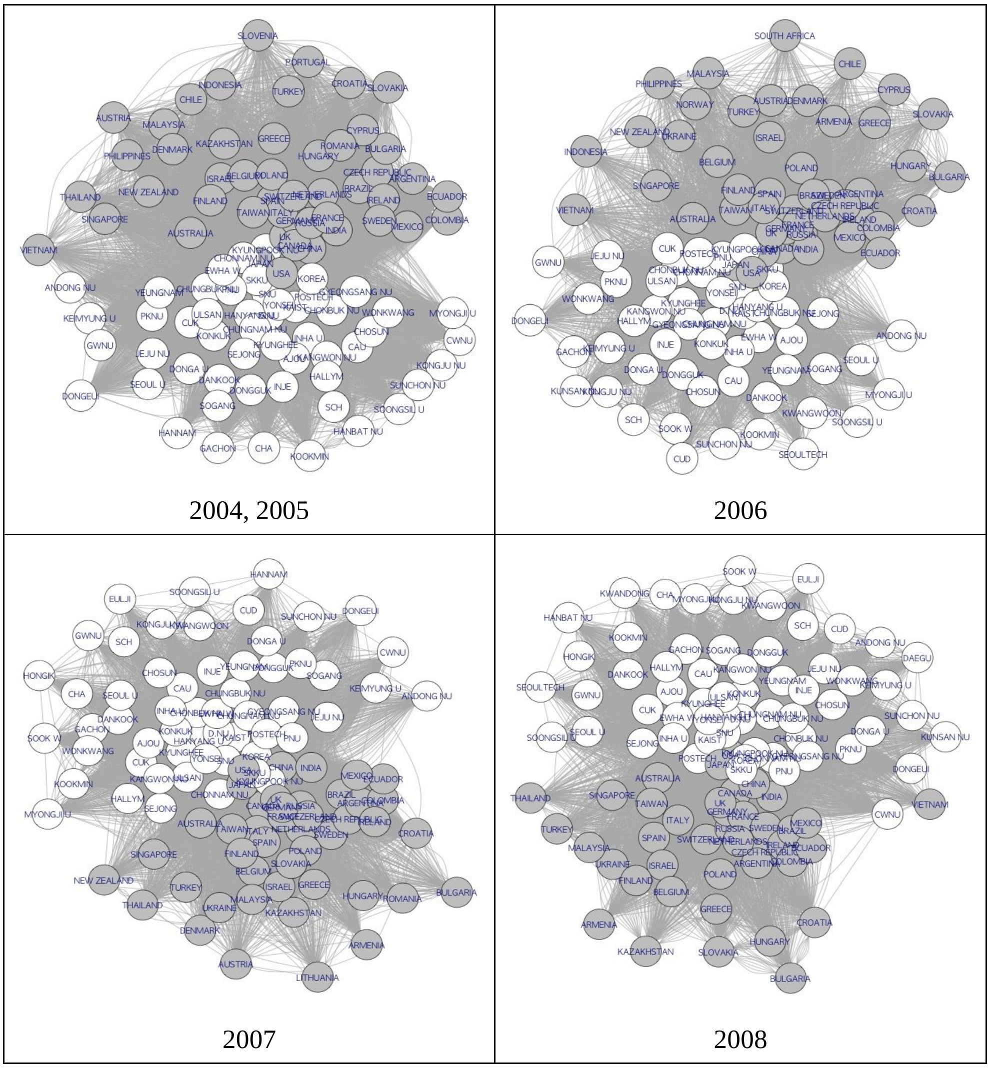


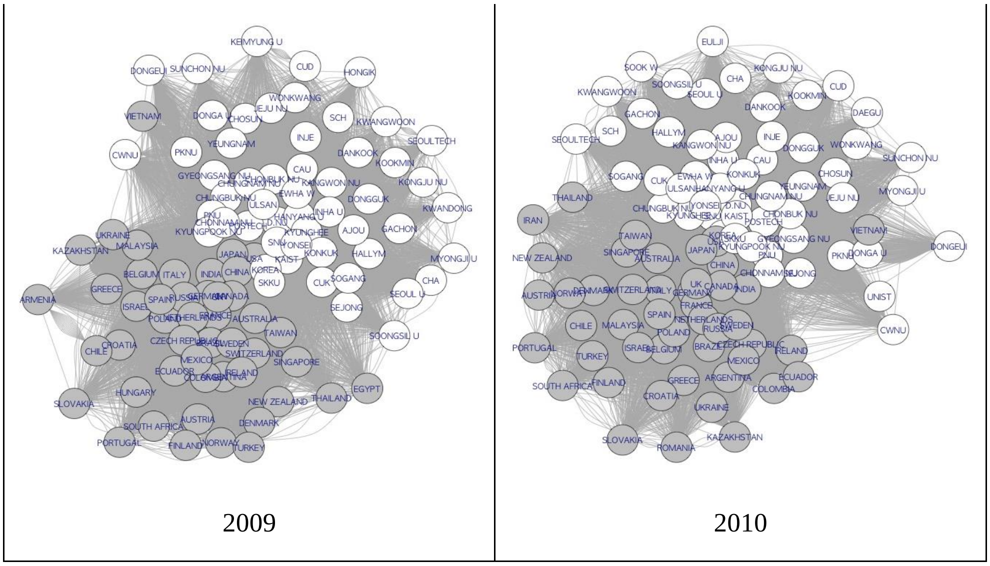
Figure 8 Scientific collaboration network, 2010
Note: Grey denotes foreign countries

The whole network is very complicated and can make it difficult to understand the structure of scientific collaboration. As there are hundreds of Korean universities and scores of other actors that are not Korean universities in the scientific community, the whole structure for all organizations in a single picture is not easy to recognize or analyze (Figure 8). Therefore, we selected the top 100 vertices for each period to investigate the patterns of scientific collaboration by Korean universities. Each panel of Figure 9 shows the network for a specific period, and it has 100 nodes with the largest value of degree centrality.

Figure 9 Collaboration network for Korean universities, 1981–2010




2004, 2005
2006
2007
2008



Note: White represents Korean universities; grey denotes both foreign countries and domestic non-university organizations.

In that figure, we can see a somewhat clear division of the collaboration network into two different spheres: The “university world” and the “external world[15].” This disunion is not a passing phenomenon, continuing over the whole study period of 1981–2010.

[15] Note that we do not refer this as “foreign world,” as non-university partners include both domestic scientific (or industrial) organizations and foreign scientists.

4.4 Structure of Scientific Network

In this section, we shall draw the network in 2006–2010 as an example of collaborative behavior.

4.4.1 The divided network

As we can see in Figure 10, although the whole network is complex, we might recognize the divided network by its color that indicates whether the node belongs to the "university world."

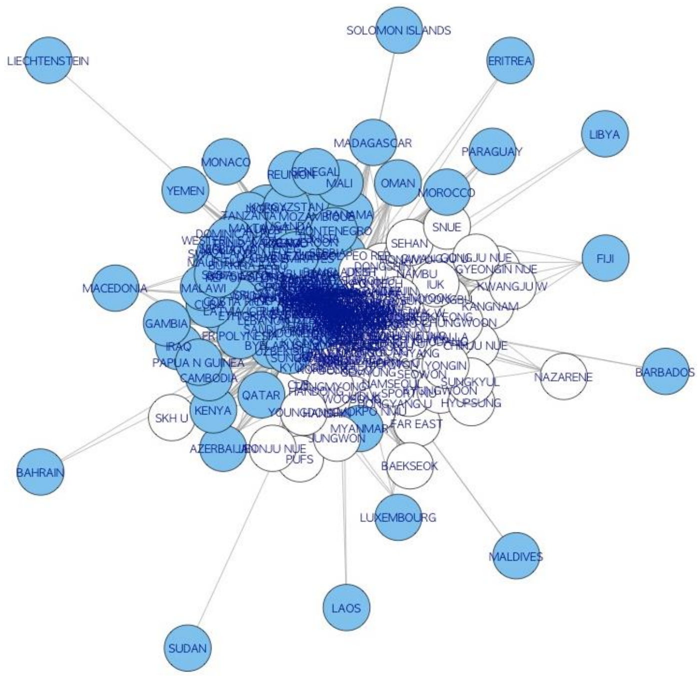


Figure 10 Scientific network during 2006–2010

Note: Sky blue denotes the external world.

4.4.2 Core/periphery structure

The "D.NU" and the "USA" are obvious "core" vertices. The universities in Figure 11 are ones who are closely related with these cores. The cores D.NU and USA formed 19.48% (24,391) and 14.29% (17,889) of the 125,200 NCR journal articles from our sample.

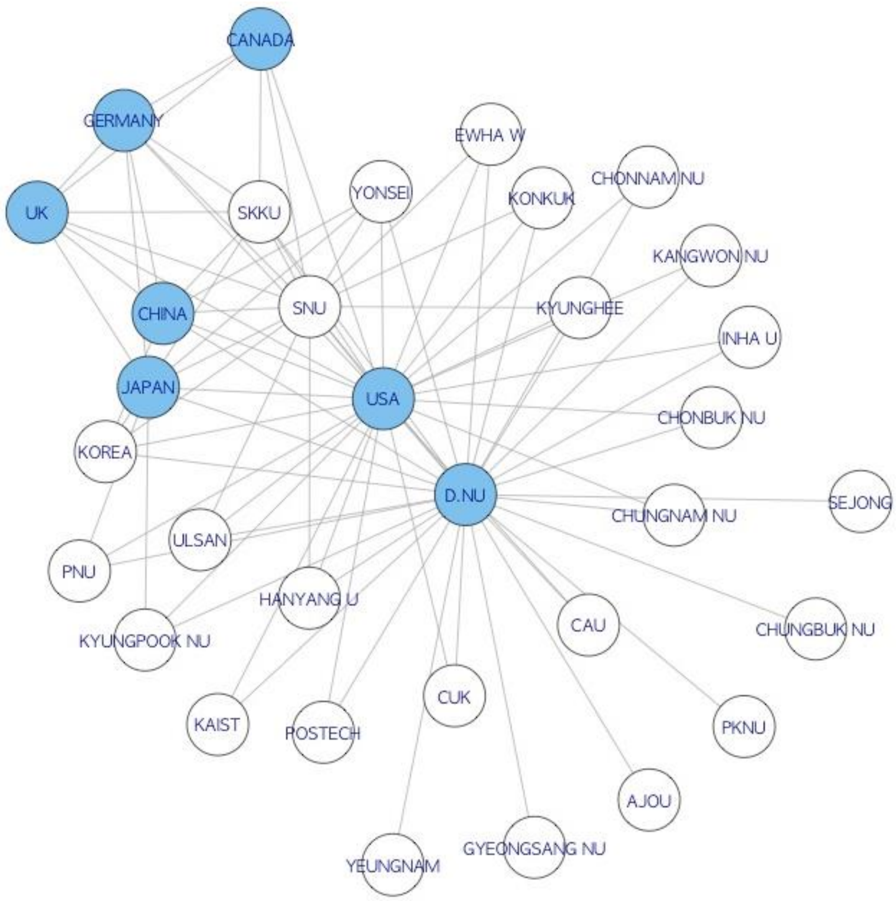


Figure 11 The scientific network constructed from selected edges (> 300) for 2006–2010.

Note: Sky blue denotes the external world.

Some vertices are only connected with the core D.NU, as shown in Figure 12. We regard these universities as being dependent on domestic-non-university organizations to produce scientific achievements: AJOU, CHUNGBUK NU, GYEONGSANG NU, PKNU, SEJONG, and YEUNGNAM.

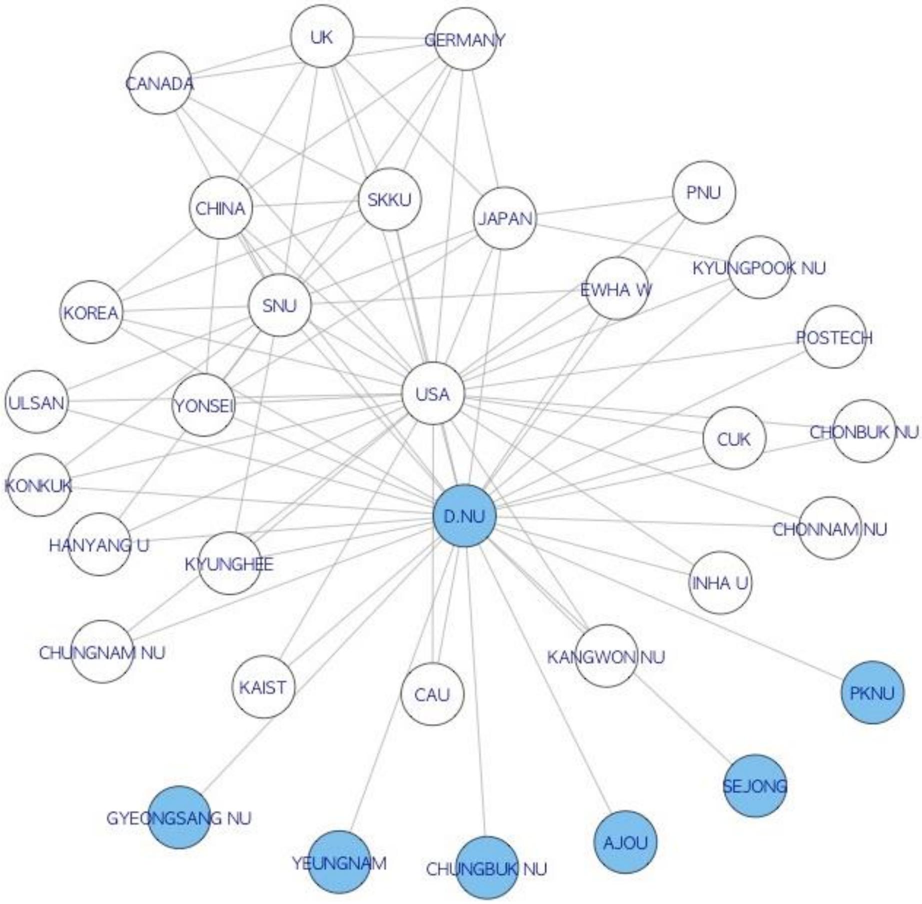


Figure 12 D.NU and some of its dependents (sky blue nodes)
Note : Edges represent 300 or more relations.

Some vertices are connected with both cores, D.NU and USA, and are not related to others. We regard these universities as ones that focus on American organizations to produce scientific achievements: CAU, CHONBUK NU, CHONNAM NU, CHUNGNAM NU, CUK, INHA U, KAIST, KANGWON NU, and POSTECH (the sky blue nodes in Figure 13).

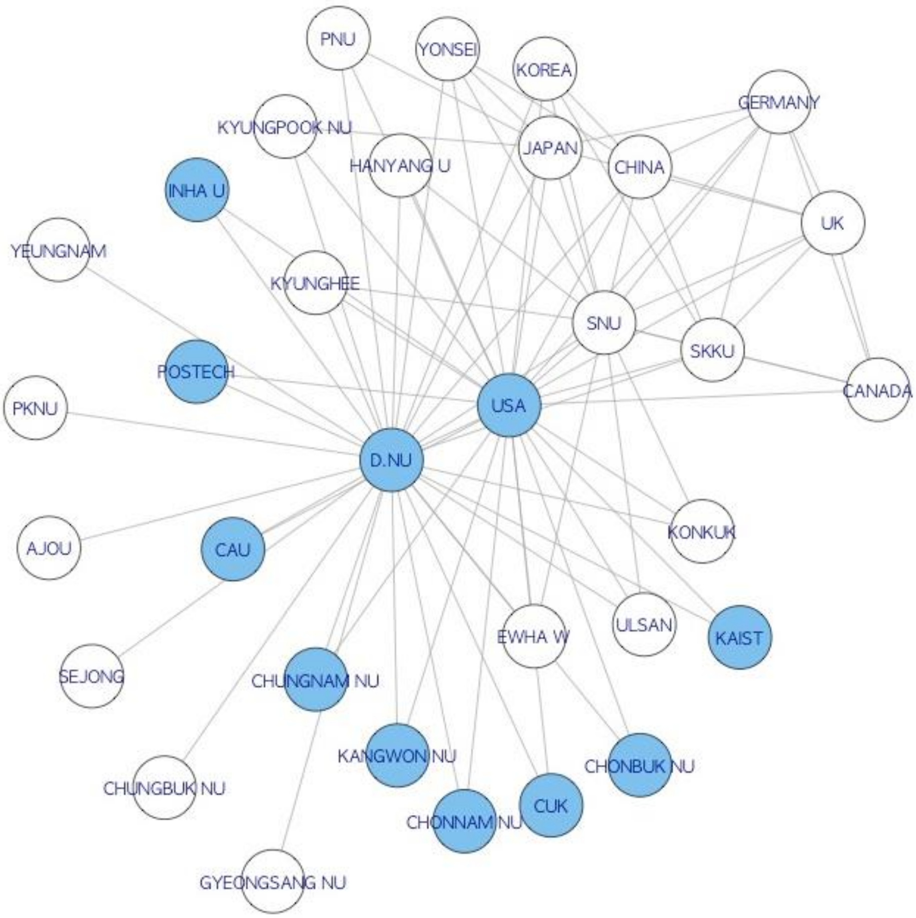


Figure 13 Universities focusing on D.NU and USA (sky blue nodes)
Note: Edges represent 300 or more relations.

To investigate the structure of the co-authorship network for Korean universities further, we removed the two cores, D.NU and USA. Now we can see that JAPAN would be removed from Figure 14, as its two vertices are no longer shown. We regard these two as universities that focus on Japanese organizations to produce scientific achievements: KYUNGPOOK NU and PNU. Note that these universities are geometrically close to Japan. JAPAN is also connected with SNU and YONSEI. The core of Figure 14 is obviously node SNU, and it has some neighbors concentrating on it: EWHA W, HANYANG U, KONKUK, KYUNGHEE, and ULSAN. CHINA has four links to this network: KOREA, SKKU, SNU, and YONSEI. SKKU and SNU have another relation with CANADA, GERMANY, and UK.

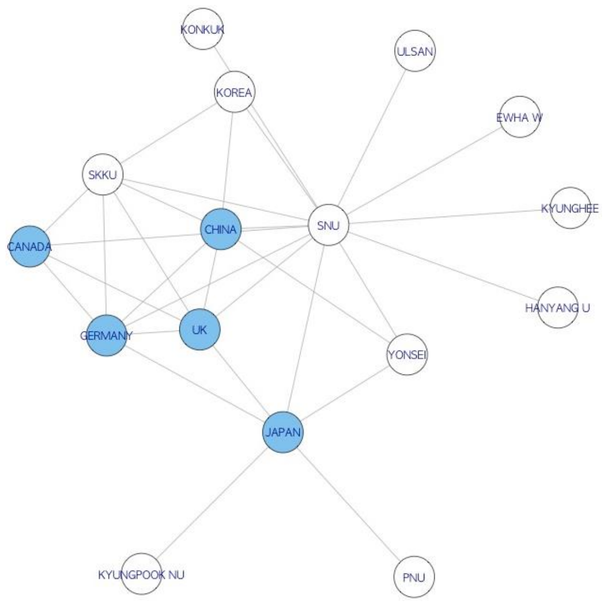


Figure 14 After removing D.NU and USA (> 300)
Note: Sky blue denotes foreign countries.

After repeating this exclusion of external-world vertices, we obtained a network independent from the external world. Note that SNU is centered and all these universities are located in Seoul. A kind of star structure is apparent, and this indicates some lack of efficiency in inter-university collaboration.

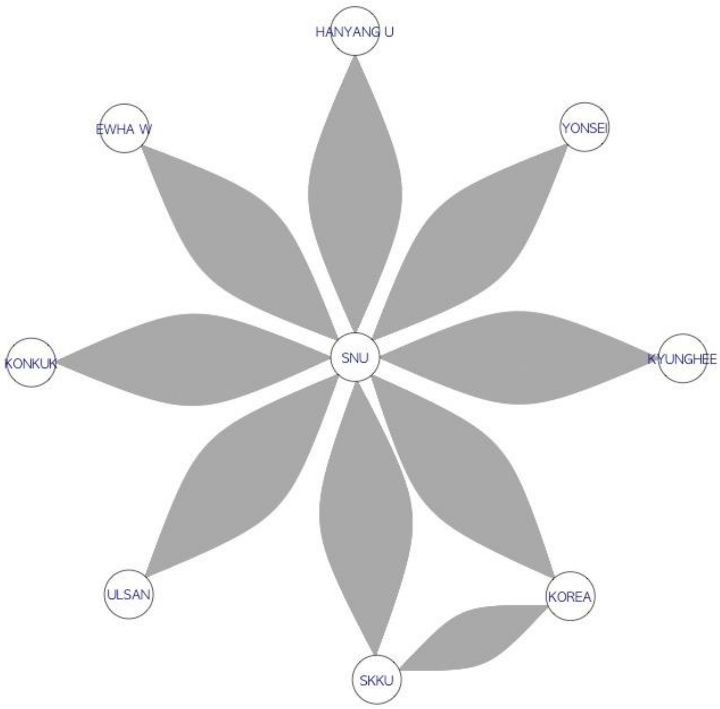


Figure 15 Universities independent from the external world (> 300).

4.4.3 Core-selection preference

Given the cores in the Korean universities' scientific network (D.NU and USA), we defined universities' "core selection preferences." The percentage of papers collaborating with a specific core is our definition, as below:

$$P_i^c = \frac{J_{ic}}{J_i} \text{ (} c \text{ is D.NU or USA)} \tag{4.1}$$

where J_{ih} is the number of journal articles having authors with the address of a university i and core c, and J_i is the number of journal articles with at least one address for university i.

Therefore, P_i^{DNU} and P_i^{USA} are the core-selection preference of university i for cores D.NU and USA, respectively.

4.4.4 Stereotyping universities

The first type is universities that are intensively collaborative with domestic-non-university (D.NU) organizations and make use of scientific strengths. Some vertices are connected with only a core D.NU, as shown in Figure 12. We regard these universities as dependent on domestic-non-university organizations to produce scientific achievements: AJOU, CHUNGBUK NU, GYEONGSANG NU, PKNU, SEJONG, and YEUNGNAM. We named this type, “D.NU-collaborative” universities.

The second type consists of universities that are intensively collaborative with the “external world,” of both cores, making use of its scientific strength. Some vertices are connected with both cores, D.NU and USA, and are not related to the others. We regard these universities as focusing on American organizations to produce scientific achievements: CAU, CHONBUK NU, CHONNAM NU, CHUNGNAM NU, CUK, INHA U, KAIST, KANGWON NU, and POSTECH. We named this type the “external-cores-collaborative” universities.

The third type is universities that are interested in collaborating with the external world further than D.NU and USA. The nodes have a third edge linked to the external world other than domestic non-universities and US organizations, such as KYUNGPOOK NU and PNU in Figure 14. We named this type the “external-collaborative” universities.

The fourth is a type that is dependent on all cores from both the “university world” and “external world” to make up for its weak scientific independence. The neighbors of SNU in Figure 15 are categorized as this type, which we named “all-cores-collaborative” universities: EWHA W, HANYANG U, KONKUK, KYUNGHEE, and ULSAN.

The fifth type is universities that collaborate with other external-world nodes in addition to D.NU, USA, and SNU. These universities are designated to the KOREA, SKKU, and YONSEI nodes, and we defined this type as “all-the-world-collaborative” universities.

All these types are listed in Table 6 below.

Table 6 Types of university in terms of their partner selection

Type	Representatives	P(D.NU)	P(USA)	P(F)	P(SNU)	P(JP)
D.NU	AJOU, CHUNGBUK NU, GYEONGSANG NU, PKNU, SEJONG, YEUNGNAM	0.212	0.114	0.181	0.082	0.042
D.NU & USA	CAU, CHONBUK NU, CHONNAM NU, CHUNGNAM NU, CUK, INHA U, KAIST, KANGWON NU, POSTECH	0.216	0.126	0.150	0.091	0.042
D.NU, USA, & (F)	KYUNGPOOK NU (JP), PNU (JP)	0.200	0.165	0.208	0.089	0.090
D.NU & USA / SNU	EWHA W, HANYANG U, KONKUK, KYUNGHEE, ULSAN	0.189	0.135	0.131	0.119	0.037
D.NU, USA, & (F) / SNU	KOREA (CN), SKKU (CA, CN, DE, UK), YONSEI (CN, JP)	0.212	0.177	0.158	0.108	0.058
Total sample		0.206	0.124	0.125	0.088	0.037

Type: External world/university world. F: Foreign country other than the USA. CN: CHINA, JP: JAPAN, CA: CANADA; DE: GERMANY

4.5 Relating Node Characteristics to Scientific Production

As we examined in the prior section, Korean universities' scientific collaboration network has special features. Given this, we shall present some network measures and closeness-to-core measures, then their impact on scientific output.

4.5.1 Centrality and scientific production

We can observe an increase in degree centralities for the university world and external world in Figure 16. Even though our data is organized based on universities, the degree of the external world is larger than the degree of university world. This indicates that there is some room for collaboration among Korean universities.

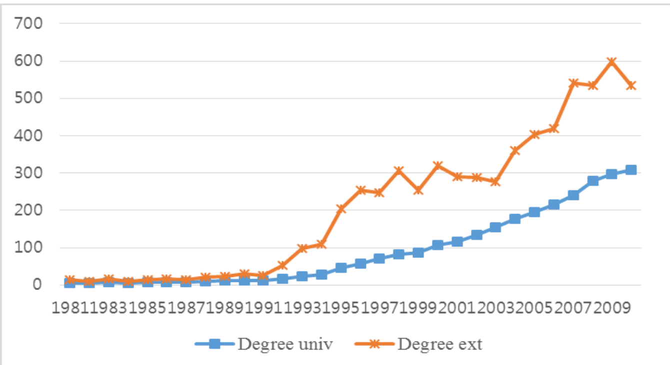


Figure 16 Degree centrality of the scientific collaboration network
Note: "univ" denotes the university world; "ext" denotes the external world.

Over time, as more connections have come to the network, the value of betweenness has decreased (Figure 17). The units for two worlds are different from one another, so it is inappropriate to compare the magnitudes of betweenness of the external world to that of the university world.

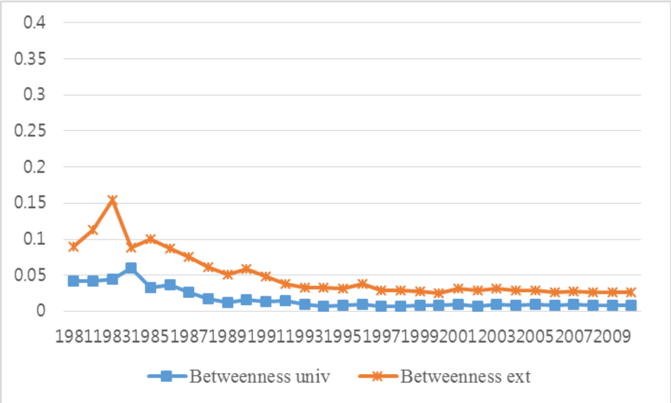


Figure 17 Betweenness centrality of the scientific collaboration network
Note: "univ" represents the university world, while "ext" represents the external world; normalized to maximum.

The more connections among nodes have come to the network over time, the more the values for closeness have decreased in Figure 18. Units for the two worlds are different from each other, it is inappropriate to compare the magnitudes of closeness from external world with those from the university world. However, it seems intuitive that the university world lacks some relative closeness.

Figure 18 Closeness centrality of the scientific collaboration network

Note: "univ" represents the university world, while "ext" represents the external world; normalized to maximum.

Table 7 Centrality scores for academic network

	Degree		Betweenness		Closeness	
	univ	ext	univ	ext	univ	ext
1981	5.85	14.882	0.042	0.089	0.744	0.840
1982	4.893	10.25	0.042	0.114	0.796	0.766
1983	7.379	16.4	0.044	0.154	0.617	0.679
1984	5.714	9.238	0.060	0.089	0.771	0.787
1985	6.689	14.529	0.033	0.100	0.772	0.774
1986	8.32	16.125	0.036	0.087	0.614	0.663
1987	8.232	14.15	0.027	0.075	0.616	0.634
1988	9.719	20.476	0.018	0.061	0.583	0.626
1989	12.556	24.269	0.012	0.051	0.582	0.617
1990	12.866	30.348	0.016	0.059	0.767	0.799
1991	13.013	25.897	0.014	0.049	0.591	0.610
1992	17.704	52.340	0.014	0.038	0.614	0.658
1993	23.890	97.556	0.010	0.033	0.578	0.599
1994	28.784	110.326	0.007	0.033	0.572	0.598
1995	45.627	203.527	0.009	0.031	0.602	0.623
1996	57.882	254.396	0.009	0.038	0.619	0.625

1997	70.607	247.507	0.007	0.029	0.604	0.625
1998	82.829	306.563	0.007	0.029	0.592	0.597
1999	87.464	253.157	0.009	0.028	0.616	0.617
2000	106.772	320.029	0.008	0.026	0.597	0.608
2001	116.069	290.594	0.009	0.032	0.623	0.635
2002	135.290	288.782	0.008	0.029	0.613	0.601
2003	154.878	277.48	0.010	0.031	0.628	0.620
2004	177.578	360.081	0.008	0.029	0.607	0.651
2005	195.772	403.295	0.010	0.029	0.625	0.653
2006	216.007	420.204	0.009	0.026	0.601	0.644
2007	241.064	541.278	0.009	0.028	0.625	0.628
2008	278.884	533.708	0.008	0.027	0.643	0.642
2009	297.206	599.14	0.009	0.027	0.608	0.633
2010	307.780	533.952	0.008	0.026	0.621	0.639

Note: "univ" represents universities only, while "ext" represents non-university organizations and foreign countries. Betweenness and closeness are normalized to maximum.

Table 8 Descriptive statistics of centralities, 2006–2010

	Min	Median	Mean	Max	Std dev	Obs
(univ)						
deg	0	58	265.524	4345.4	552.984	141
bet	0	0.000	0.006	0.284	0.027	141
clo	0.484	0.569	0.564	0.892	0.154	141
(ext)						
deg	7	64.5	453.528	7489	1077.644	115
bet	0	0.000	0.023	0.986	0.126	115
clo	0.530	0.603	0.618	1	0.089	115

Note: "univ" represents universities only, while "ext" represents non-university organizations and foreign countries. Betweenness and closeness are normalized to maximum

Table 7 provides some detailed results for network centralities in figures 16–18. We might observe the details of increases in the degree of centrality and of the decrease in betweenness and closeness centralities.

Then we might observe descriptive statistics of centralities for the period 2006–2010. In that period, there are 115 vertices belonging to the external world. Let us note that the core/periphery structure of Korean universities' collaboration networks leads to centering the distributions of centralities. All these centrality variables are positively related with the scientific production of a university (Table 9). In spite of the high correlation among them, the inability to discriminate caused by their great dispersion made us think that network centralities would not have an impact on scientific production.

Table 9 Correlation among network centralities and scientific production in the world of universities

	sci	deg	bet
deg	0.959***		
bet	0.788***	0.871***	
clo	0.872***	0.829***	0.574***

Significance level: *** p<0.01

4.5.2 Core selection and scientific production

In the case of the core selection preference for universities, distributions have lesser variances compared to network centralities (Table 10). For the average university in Korea, the share of collaboration including D.NU (USA) is recorded as 0.204 (0.125), indicating one of every five (eight) publications is co-authored with domestic-non-university (US) organizations.

Table 10 Descriptive statistics of partner preference for Korean universities, 2006–2010

	Min	Median	Mean	Max	Std dev	Obs
Pdnu	0	0.202	0.204	0.444	0.083	141
Pus	0.015	0.112	0.125	0.476	0.073	141

While dispersion conditions are good for partner preferences, the correlation with scientific production seems not; the two partner preferences are negatively correlated with each other (Table 11).

Table 11 Correlation among scientific production and partner preference

	articles	Pdnu
Pdnu	0.024	
Pus	0.096	-0.178**

Significance level ** p<0.05

4.6 Conclusion: Partner selection strategy for more scientific production

This chapter aims to supplement Chapter 3 with a bigger picture of the detailed structure of collaboration tendencies from Korean university academia. We took social network analysis (SNA), which appears to be appropriate for examining the interactive relationships such as co-authorships, to depict structural relations. For this purpose, we defined the “external world” as the coupling of domestic non-university institutions with organizations from foreign countries, discriminating them from the “university world.” Then we drew bigger pictures to investigate the few structural relationships in a scientific co-authorship network for Korean academia, which consists of 141 universities and 115 external units.

The results imply that a scientific collaboration network within Korean university spheres presents three characteristics: (a) A divided network that indicates less collaboration beyond national borders; (b) A core/peripheral network that reveals the huge importance of a few cores; (c) A star network that demonstrates the fragmentary nature of collaboration among Korean universities.

From a strategy perspective, our results highlight an important phenomenon. Universities are required to conduct more collaborative research with cores inside their networks when they need to increase their scientific production. In the core/periphery structure, there is no special space other than the center. Therefore, the closeness to cores would decide the power of a given actor. In the case of our study, USA seems a powerful option.

From a policy perspective, the results also highlight two relevant circumstances; first, the structure of collaboration at the “university world” level is as star-shaped network, centered on SNU. This structure is very inefficient for disseminating information to the whole network. The fragmented collaboration also appeared as the inter-tier tendencies of Korean universities in scientific partner selection (Ahn et al., 2014), and became visible as a star network after removing external world nodes in this chapter. It is required that Korean universities recognize other universities in their world as potential partners rather than as competitors, and that policy makers recognize this phenomenon as a problem we should solve to improve the efficiency of our scientific system at the “university level.” Second, the USA is the core of the knowledge network for Korean universities. If we want a situation in which one or more universities are the core of the network, we need more collaboration between universities.

Chapter 5. Knowledge Production Function for Korean University

5.1 Introduction

Academics and politicians increasingly admit that scientific research is very important to provide the foundations for both innovation and competitiveness (Crespi, 2007). The global world are competitively increasing national research expenditure in the Korea and elsewhere. There is, however, little systematic evidence on how R&D investments can generate an increase in levels of scientific output, an advance in innovative output such as patent, economic growth, and conclusively wealth achievement (Crespi & Geuna, 2008). Many of existing studies are interested in the effects of scientific research on the firms' innovation (Jaffe, 1989; Mansfield, 1991; Narin et al., 1997). Much of literature concentrates on the contribution of basic research to productivity growth (Adams, 1990). Furthermore, many recent studies have analyzed university-industry relationships including technology transfer and intellectual property rights (for an overview see Organisation for Economic Co-operation and Development [OECD], 2002). In the literature on university-industry, a considerable body of work focus on the relevance of contract research, collaborative R&D, consulting, and informal relationships for knowledge (or technology) transfer between university and industry

(Perkmann et al., 2013). The contribution of public R&D investments is likely to be examined by analyses of national productivity gains or economic growth (Guellec & van Pottelsberghe de la Potterie, 2001; Hall et al., 2010; OECD, 2003). However, only a very few works have concentrated on the relating R&D investment to the measure of scientific outputs (of particular focus are Adams & Griliches, 1998; Crespi, 2007; Crespi & Geuna, 2008).

The aim of this chapter is to develop a model of knowledge production function embedding a better understanding of the collaborative characteristics in scientific production process. The analysis is undertaken from a 'university world' viewpoint, based on a hypothesis that university should be affected by both an 'inside' and an 'outside' of the university world. We employed a sample of 144 Korean university schools for which we have reliable information on R&D expenditure and scientific collaboration. As a proxy for output we used publications from total scientific journals and from leading journals. Most inputs and outputs for this sample of universities are from the "*Higher Education in Korea" service*[16] (hereafter HEiK), provided by Korean Council for University Education, which allows the user to search information recorded during 2008–2012. For the scientific collaboration, which is measured as co-authorship in the articles, the *1981–2010 South*

[16] Provisions in 6th section of the *Act on Information Disclosure of Educational Institutions* has obligated the educational institutions to publicly disclose information and to be regulated relevant details. The Web site address for the *Higher Education in Korea service* is www.academyinfo.go.kr.

Korea NCR[17] (hereafter NCR), which is used for the earlier chapter, is utilized for collaboration characteristics of each university during 2006–2010. Given these data, this chapter provides an assessment of the returns to investment in university science, particularly considering the existence and characteristics of two spillovers: the one is ‘cross-university’ and the other is ‘from US.’ The results of this chapter implicate that spillover effect has a real existence and a deep relation with scientific collaboration, while this effect is somewhat weak. The small spillover effect request either much more collaboration or much efficient co-operation.

The remainder of this chapter is organized as follows. Section 2 introduces the data set we used for analysis and describes it. Section 3 presents the models and estimation strategy. The empirical results are presented and discussed in Section 4. The last section concludes and suggests for further study.

[17] NCR is an abbreviation for ‘national citation report.’

5.2 Data

In this section, we describe the structure and source of our dataset, variables comprising knowledge production function, and descriptive statistics for variables used.

5.2.1 Data sources

The sources of data fall into two groups depending on whether they are related to personnel and financial information or whether they refer to the inter-university collaboration.

The first one, personnel and financial information, is from website named "*Higher Education in Korea" service* (HEiK), provided by Korean Council for University Education, which allows the user to search information on Korean universities. From this source, we excavated data about the university system by institution for 2008–2012 period.

The other source of our data is the *1981–2010 South Korea NCR* (NCR), which is used for the earlier chapter. This source, as a subset of the *Web of Science* (WoS) database, includes bibliographic and citation information on 297,658 regular scientific journal articles that have at least one Korean address among the addresses for authors. From this source, we excavated data about the collaboration structure for 1981–2010 period.

5.2.2 Variables

Here are variables for building a basic model of the knowledge production function approach.

5.2.2.1 Research output

There are several problems in measuring the research-output variables, the dependent variables in this study. The scientific production process might yield three categories of research output: new knowledge, human resources, and new technologies (Crespi & Geuna, 2008). In this study, we have focused on the first category of research output. Though there are no direct measures for the new knowledge, publications and citations are commonly used as proxies. We used two measures for research output variable which are available in our data. The one is the 'total-journal' articles and the other is the 'leading-journal' articles. The total-journal articles indicates the number of articles from all sorts of scientific journals, and the leading-journal articles stands for the number of articles from selected journals as SCIE, SSCI, A&HCI, and SCOPUS. The database, HEiK, provides these scientific achievement measures by university. It is note that, however, this figures are adjusted with contribution of author and the total number of authors in a paper as follows:

$$Q_l = \begin{cases} 1 & \text{(sole-auhored)} \\ \dfrac{2}{N+2} & \text{(first/corresponding)} \\ \dfrac{1}{N+2} & \text{(otherwise)} \end{cases} \tag{5.1}$$

where $0 < Q_l \leq 1$ is research output and N is the total number of authors in a scientific journal article, for a full-time faculty member l. Then this values for a faculty members are aggregated to university level, Q_i, for a university i.

$$Q_i = \sum_{l \in i} Q_l \tag{5.2}$$

We shall notice that this adjusted measure is largely affected by the number of full-time faculty members.

As typically publication is measured as count data, many studies (e.g. Crespi, 2007) estimated models using Poisson regression or negative binomial one[18]. On the other hand, the dependent variable for this study is not a typical count variable, therefore we could not use Poisson and negative binomial distributions for modelling.

---

[18] The standard Poison regression assumes that a conditional variance and a conditional mean of the dependent variable. The negative binomial distribution, on the other hand, allows for consistent and efficient results when the dependent variable is a count with overdispersion. Here overdispersion means that the dependent variable’s variance is larger than its mean. See Greene (2012).

5.2.2.2 Research input

"R&D expenditures may differ in type but their object is always to increase the stock of knowledge in order to find new applications and innovations." — Hall et al. (2010, p. 1035)

R&D expenditures by university are also retrieved from HEiK database. The dataset provides an information on the funding sources together: from the government, private sector, university's own invest, and foreign organizations. The descriptive statistics for this variable are on the table 12.

Table 12 Research expenditure at Korean universities, 2008–2012 (1,000 KRW, 2005 constant prices)

	Min	Median	Mean	Max	Std dev	Obs
(2008)						
R	0	7198380	23039451	369940134	45288145	144
Rgov	0	5676876	17534358	326162975	36395516	144
Rpriv	0	757950	2895338	42252427	6506596	144
Runiv	0	932188	2524939	38335085	4957739	144
Rintl	0	0	84816	2489990	321926	144
(2009)						
R	46083	8534613	25704059	396352552	48669185	144
Rgov	27650	6215906	20127320	364606135	40713438	144
Rpriv	0	814144	2874762	35347540	5740520	144
Runiv	0	914187	2639356	30204001	4686668	144
Rintl	0	0	62621	1804939	232213	144

(2010)						
R	55249	8187396	27470663	437028599	54040496	144
Rgov	28470	6095521	21432972	400108282	44797727	144
Rpriv	0	894583	3333213	43376959	7043950	144
Runiv	0	830774	2641572	34409630	4677467	144
Rintl	0	0	62906	1423854	217645	144
(2011)						
R	282420	7968566	28731702	447169620	55704548	144
Rgov	75951	6483737	22578529	398662987	45592050	144
Rpriv	0	852048	3475819	46683685	7807790	144
Runiv	0	929766	2600581	25496612	4436607	144
Rintl	0	0	76773	3162750	330360	144
(2012)						
R	83393	7677225	29449748	412199032	56215982	144
Rgov	83393	6281491	23094243	396845091	46817833	144
Rpriv	0	693901	3574372	46890905	7773073	144
Runiv	0	837118	2695571	36998535	4875259	144
Rintl	0	0	85562	1740975	281850	144

Note: R is total ressearch expenditure

Rgov is Research expenditure funded from government

Rpriv is Research expenditure funded from private sector

Runiv is Research expenditure funded by university itself

Rintl is Research expenditure funded from foreign country

In case of the knowledge stock, which is generally a function of the R&D expenditure, there is a great controversial issue on the specification of it. Griliches (1979) states three major issues considering the measurement of knowledge capital: (a) the fact that a research process takes time and that current investment may not have an immediate effect on productivity until several years; (b) past R&D expenditures depreciate and become outmoded; (c) that the knowledge level for a given research unit is not only determined by 'own' investments, but is also influenced by others' knowledge, through knowledge spillovers.

Traditionally, the influence of past R&D expenditures has been modelled using a linear inverted U- or V-shaped function (see e.g. from Adams & Griliches, 1998; to Crespi, 2007).

$$K_t = \begin{cases} \sum_{j=0}^{J} W_j R_{t-j} \\ \prod_{j=0}^{J} R_{t-j}^{W_j} \end{cases} \tag{5.3}$$

5.2.2.3 Variables controlling for university-specific characteristics

The number of full-time faculty members is a main control variable in this study as in others (e.g. Crespi, 2007). As mentioned above, our measure of scientific output of a university is summation of adjusted research outputs by each full-time faculty member. This situation automatically gives the faculty size a great power to output measure.

The number of students is also a major variable controlling for the effect of research-teaching relationship. Studies in higher education (HE) has been interested in the phrase 'research/teaching nexus' referring to the relationships between research and teaching (for an overview see Jenkins, 2004; Marsh & Hattie, 2002). Most works in the literature have paid attention to the role of research or research skills in relation to student experience in HE (Fairweather, 2005; Ramsden & Moses, 1992). We used the numbers of both graduate and undergraduate students as measures for this control variable.

In Korea, there are several specialized universities focusing to one or some particular disciplines of science. We used dummies for these specialized schools: ‘tech’ is the one for the universities specializing into science & technology (including engineering), and ‘edu’ is the one for institutions providing training for future public elementary school teachers. Note that theses specialized organizations are operated by the government.

Table 13 Variables for knowledge production function

Variable	Definition/Description
(output)	
arti	Number of scientific journal article Sum of personal achievements (author-size adjusted paper counts) for all full-time faculty members
sci	number of global-leading scientific journal article same as 'arti' but only journal belonging to SCI(E), SSCI, A&HCI, or SCOPUS
(input)	
R	Research expenditure by university can be divided by funding source (government, private sector, university itself, foreign country)
K	Knowledge Capital/Stock Function of past 'R's, also divided by funding source like 'R'
(control)	
fac	Number of full-time faculty members This is a control variable not an independent one; it is because 1) dependent variable is intermediate and 2) we are not interested in adjusting faculty size of schools
und	Number of enrolled undergraduate students Similar to 'fac'
post	Number of enrolled graduate students Similar to 'fac' and 'und'
tech	Dummy for indicating universities which concentrate science and technology discipline
edu	Dummy for indicating university of education which provide training for future public elementary-school teachers

5.2.3 Descriptive statistics at university level

Our analysis is based on 144 universities in Korea. We excluded the remaining university schools by reason of incomplete information either on university itself (from HEiK) or scientific collaboration (from NCR). As we see in table 14, the scientific production of Korean universities has grown from 2011 to 2012, as the knowledge capital stock. The numbers for faculty size and students have not change, relatively. Though not strongly, this supports our selection of dependent, explanatory, and control variables.

Table 14 Descriptive statistics: research outputs, knowledge stock, and control variables

	Min	Median	Mean	Max	Std dev	Obs
(2011)						
sci	1	37.1	141.3	1,786.2	257.5	144
arti	31.2	264.8	457.3	2,937.6	495.1	144
K	161,487	7,872,213	25,479,558	399,918,459	49,024,414	144
fac	55	311	453	2,164	386	144
post	37	1,102	2,565	18,122	3,353	144
und	891	8,298	10,166	27,229	6,518	144
(2012)						
sci	1.6	45.5	164.8	1,935.0	287.8	144
arti	46.2	283.3	484.3	3,006.7	514.2	144
K	168,767	7,918,046	27,344,272	429,394,842	53,035,767	144
fac	57	311	464	2,178	394	144
post	31	1,057	2,565	17,906	3,349	144
und	1,256	8,331	10,018	27,212	6,677	144

K is in 1,000 constant 2005 KRW

5.3 Methodology

In this section, we describe the models and estimation strategy for developing knowledge production function for Korean university considering scientific collaboration effect.

5.3.1 Modelling framework

Here we present details of modelling the scientific production process at universities in Korea based on the knowledge production function.

5.3.1.1 Science production function

Like many studies estimating knowledge production function, we begin our estimation with Cobb-Douglass knowledge production function from Griliches (1979).

$$Q_{it} = A_i K_{it}^{\beta} e^{\lambda_t + u_{it}} \tag{5.4}$$

It is used widely relating knowledge outputs to knowledge capital (a. k. a. knowledge stock). Q is some measure of intermediate output, K is a measure of the current state of scientific knowledge, and u represents all other unmeasured determinants after the addition of a time trend to indicate the systematic component of the unmeasured factors, such as spillovers. Note that the coefficient β measures the elasticity of output with respect to own knowledge capital.

For ease of use, equation (5.1) is likely to be transformed to the log-linear version, as below:

$$q_{it} = c + \beta k_{it} + \gamma X_{it} + \lambda_t + \omega_i + u_{it} \tag{5.5}$$

In deriving this, we assumed implicitly that the logarithm of technical progress A can be written as the sum of a time effect λ_t and a university-specific effect ω_i

In case of the knowledge stock, there is a great controversial issue on the specification of it. Griliches (1979) states three major issues considering the measurement of knowledge capital: (a) the fact that a research process takes time and that current investment may not have an immediate effect on productivity until several years; (b) past R&D expenditures depreciate and become outmoded; (c) that the knowledge level for a given research unit is not only determined by 'own' investments, but is also influenced by others' knowledge, through knowledge spillovers.

Traditionally, the influence of past R&D expenditures has been modelled using a linear inverted U- or V-shaped function (see e.g. from Adams &

Griliches, 1998; to Crespi, 2007). As our dataset has information on R&D expenditure for only 5 years(2008–2012), we selected three-year lags being weighted by an inverted V-shaped function.

$$K_t = \frac{1}{4} R_{t-1} + \frac{1}{2} R_{t-2} + \frac{1}{4} R_{t-3} \tag{5.6}$$

5.3.1.2 Premium (or discount) for each funding source

We adopted specification of Crespi (2007) to further augment basic model by accounting for the separated effects of various funding sources on scientific production: university itself, government, private sector, and foreign country.

$$\begin{aligned} q_{it} &= c + \gamma X_{it} + \lambda_t + \omega_i + u_{it} + \\ &\beta \ln[K_{it}^{G} + (1+\Omega_1)K_{it}^{U} + (1+\Omega_2)K_{it}^{P} + (1+\Omega_3)K_{it}^{F}] \end{aligned} \tag{5.7}$$

The below term of equation (5.4) can be transformed to sum of knowledge capital and capital shares from funding sources.

$$
\begin{aligned}
&\beta \ln[K_{it}^{G} + (1+\Omega_1)K_{it}^{U} + (1+\Omega_2)K_{it}^{P} + (1+\Omega_3)K_{it}^{F}] \\
&= \beta \ln[K_{it}(1+\Omega_1 \frac{K_{it}^{U}}{K_{it}} + \Omega_2 \frac{K_{it}^{P}}{K_{it}} + \Omega_3 \frac{K_{it}^{F}}{K_{it}})] \\
&(\because K^{G} + K^{U} + K^{P} + K^{F} = K) \\
&= \beta k_{it} + \beta\Omega_1 \frac{K_{it}^{U}}{K_{it}} + \beta\Omega_2 \frac{K_{it}^{P}}{K_{it}} + \beta\Omega_3 \frac{K_{it}^{F}}{K_{it}} \\
&[\because \ln(1+z) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{z^n}{n} \approx z \text{ for small } z]
\end{aligned} \tag{5.8}
$$

After transforming, we can get the model having total knowledge capital and proportions of each funding source to it.

$$
\begin{aligned}
q_{it} = c + \gamma X_{it} + \lambda_t + \omega_i + u_{it} + \\
\beta k_{it} + \beta\Omega_1 \frac{K_{it}^{U}}{K_{it}} + \beta\Omega_2 \frac{K_{it}^{P}}{K_{it}} + \beta\Omega_3 \frac{K_{it}^{F}}{K_{it}}
\end{aligned} \tag{5.9}
$$

5.3.1.3 Cross-university knowledge spillovers

We defined the ‘cross-university K’ linked to each school as a sum of the weighted knowledge stocks of all the other universities. We have to reveal that this definition is a minor alteration of “international R&D” by Crespi and Geuna (2008). The variable is defined as equation (5.10), where w_{li} is knowledge proximity between university l and i.

$$CU_{it} = \sum_{l \neq i} w_{li} K_{lt} \tag{5.10}$$

As Crespi and Geuna (2008) used information on international scientific co-authorship, here we used inter-university scientific co-authorship to build a matrix for knowledge proximity among schools. Equation (5.10) is a function for an element of the proximity matrix as follows:

$$w_{li} = \frac{C_{li}}{\sum_{j} C_{lj}} \tag{5.11}$$

where C_{li} indicates the number of co-authorship journal articles in the NCR. Therefore w_{li} is the share of university i among total co-authorships of university l. We counted and calculated those numbers by year, and then averaged numbers for making a final matrix.

Inserting CU_{it} into model (5.5) gives an extended specification for allowing the effects of knowledge spilled over from all the other universities inside the national border, as:

$$q_{it} = c + \beta_K k_{it} + \beta_{CU} cu_{it} + \gamma X_{it} + \lambda_t + \omega_i + u_{it} \tag{5.12}$$

where lower case stands for logarithm of each variable.

Although this cross-university knowledge-spillovers effect is considered significantly, there is a limit to this approach arising from the fact that scientific production has been characterized by inter-sectoral or international cooperation, it became even more relevant. We should, therefore, consider scientific collaboration with the outside of the university world, as a couple of sections below.

5.3.1.4 Network position and core preference

As we examined in Chapter 4, Korean universities are placed in a core/periphery network in terms of scientific collaboration. This situation leaded us to a question that network position or core preference of a university might also influence its scientific production. The specification of modelling is like below:

$$q_{it} = c + \beta_K k_{it} + \beta_{CU} cu_{it} + \gamma_N N_i + \gamma_X X_{it} + \lambda_t + \omega_i + u_{it} \qquad (5.13)$$

$$q_{it} = c + \beta_K k_{it} + \beta_{CU} cu_{it} + \gamma_P P_i + \gamma_X X_{it} + \lambda_t + \omega_i + u_{it} \qquad (5.14)$$

Equations (5.13) and (5.14) are similar specification for adding network position and core preference respectively, where N_i is the network position measured by node centrality and P_i is the core preference at 4.4.3 of this volume.

5.3.1.5 Knowledge spillovers from external world

Given the open character of scientific production we would predict significant interactions with external-university world, and Chapter 4 made certain of the fact that collaboration with US organizations is crucial to Korean universities' scientific production. It is therefore pertinent to examine whether the US' investment in science and research could affect knowledge production of Korean universities. For auditing this kind of phenomenon we explore whether spillovers occur between Korean universities and US organizations. 'Spillovers,' in this context, refer to that part of the increase in the scientific output of a given university as a result of the higher education research & development (HERD) investments of USA.

We defined the 'US' R&D relevant to each university as the science budget of the USA, weighted by core preference to USA (P_i^{USA}), as follows:

$$US_{it} = P_i^{US} K_{US,t} \tag{5.15}$$

where $K_{US,t}$ is the knowledge stock of the USA, a function of past HERD with same lag structure as one that Korean universities have (see equation 5.6).

Inserting US_{it} into model (5.12) gives an extended specification for allowing the effects of knowledge spilled over from both all the other universities inside the university world and the core of scientific collaboration network from external world, as:

$$q_{it} = c + \beta_K k_{it} + \beta_{CU} cu_{it} + \beta_{US} us_{it} + \gamma_X X_{it} + \lambda_t + \omega_i + u_{it} \qquad (5.16)$$

where us_{it} is the logarithm of US_{it}, such as other variables.

5.3.2 Estimation issues

Here we elucidate the process while estimating knowledge production models presented above. In a word, our sample failed the homoscedasticity assumption for multiple regression. Therefore we exploited the available remedies when heteroscedasticity occurs. To detail this process by giving a concrete example, we show a practice of estimating the basic knowledge production expressed as equation (5.5).

5.3.2.1 OLS and diagnosis

We estimated using the ordinary least squares (OLS) technique first. Then we selected regressors using the Akaike information criteria (AIC)[19]. In case of our basic model, there was no regressor which should be excluded. A diagnosis on the multicollinearity problem was also made using the variance inflation factor (VIF). Although 'faculty' variable has somewhat higher value

[19] For AIC, see Akaike (1974).

of VIF as 8.561, it is below the cutoff value of VIF[20] (10) and we concluded multicollinearity is not a problem.

5.3.2.2 Outliers and influential observations

After selecting regressors and testing multicollineartity, we searched outliers and influential observations to exclude from estimating. For outlier detection, we used Bonferroni p-values for studentized residuals (Fox, 2008). In case of basic model (equation 5.5), our sample had had three observations having Bonferroni p below 0.05 then those outliers are excluded. Moreover, Cook's distance was used to detect some influential observations (Cook, 1977). There were no observations having Cook's distance above 1 in our sample for equation 5.5.

5.3.2.3 Heteroscedasticity

As our data contains only information for the years 2011 and 2012, we pooled these two years and added a year dummy for 2012 capturing the trend for researchers to produce research outputs. The resulted cross-section data is commonly expected to have a condition termed heteroscedasticity (Johnston & DiNardo, 1997). Heteroscedasticity means a condition that the disturbance variances are not same at all sample points. The OLS technique could not lead

[20] However, this value is arbitrary and not especially helpful (Wooldridge, 2008).

to desirable properties of resulted estimators when the assumption of homoscedasticity is violated. The properties of OLS under heteroscedasticity are: (1) OLS estimator is unbiased and consistent; (2) OLS estimator is inefficient; (3) Conventional OLS test statistics are invalid because conventional standard errors for coefficients are incorrect. Generally, two solutions are used to deal with heteroscedasticity: the generalized least squares (GLS) and heteroscedasticity-robust inference (Wooldridge, 2002). GLS is used when we know the structural form of heteroscedasticity. Practically some feasible GLS estimator is computed from estimated functional form of the heteroscedasticity (Johnston & DiNardo, 1997; Wooldridge, 2003). Heteroscedasticity-robust inference uses heteroscedasticity-robust standard errors to compute heteroscedasticity-robust *t* statistics; heteroscedasticity-robust standard errors are computed from heteroscedasticity-corrected covariance matrices (a.k.a. White-Huber covariance matrices[21]). We also shall note that the Breusch-Pagan test (Breusch & Pagan, 1979) is used to examine the presence of heteroscedasticity.

In case of our estimating the basic knowledge production function, heteroscedasticity was rendered as Figure 18. We can see the variance of residuals becomes smaller as fitted values grow, and this could be explained

[21] The heteroscedasticity-robust variance matrix was introduced by White(1980). Before this, Huber (1967) discovered robust variance matrices.

by that the smaller universities have the broader spectrum of diversity. The Breusch-Pagan statistic was 41.539 and this supports rejecting homoscedasticity assumption for the OLS. Therefore we used available remedies for heteroscedasticity such as corrected OLS, WLS, and FGLS. The empirical results are presented in next section.

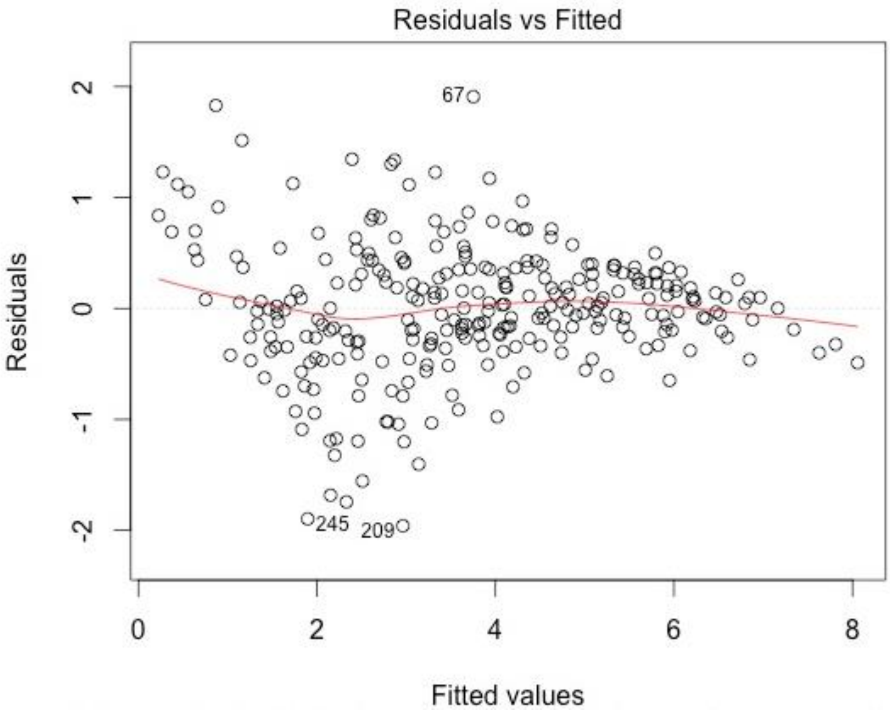


Figure 19 Scattered plot of residuals from estimating the basic knowledge production function (dependent variable log ‘sci’)

5.4 Empirical Results

Here we present results from empirical study estimating knowledge production function specified as former section. Then we discuss about result and provide some implications from variety perspectives.

5.4.1 Scientific returns to R&D

Table 15 and Table 16 show the results for the basic model of knowledge production function as equation (5.5). For both outputs, there exists positive, significant returns to R&D (0.59 and 0.10). The larger number of full-time faculty members is linked to higher scientific production. It is interesting that the number of students, both postgraduate and undergraduate, has contrary influence to two different scientific outputs: students have negative but not significant influence to ‘sci’ while they have positive and significant influence to the number of total article. The universities focusing on science and technology are good at producing science in both cases of result types. The university schools of education also have strength at scientific achievement, although it is not significant in case of leading-journal articles. The production of ‘sci’ grew in 2012 by 0.16%p, while production of ‘arti’ did not change.

Table 15 Results from basic knowledge production function (dependent variable log ‘sci’)

	OLS	corrected OLS	WLS(fac)	FGLS
K	0.557		0.579	0.593
	(0.050)***	(0.073)***	(0.049)***	(0.048)***
fac	1.335		1.283	1.210
	(0.131)***	(0.160)***	(0.121)***	(0.109)***
post	-0.126		-0.114	-0.092
	(0.069)*	(0.091)	(0.068)*	(0.064)
und	-0.097		-0.106	-0.106
	(0.107)	(0.118)	(0.100)	(0.085)
tech	0.756		0.690	0.655
	(0.186)***	(0.158)***	(0.183)***	(0.135)***
edu	0.430		0.297	0.263
	(0.231)*	(0.357)	(0.254)	(0.393)
time	0.197		0.182	0.156
	(0.070)***	(0.070)***	(0.064)***	(0.054)***
cons	-11.303		-11.333	-11.271
	(0.710)***	(0.883)***	(0.693)***	(0.663)***
Adj R-sq	0.890		0.906	0.921
Breusch-Pagan	41.539***		29.691***	16.545**

Significance levels: *** p<0.01; ** p<0.05; * p<0.1

Table 16 Results from basic knowledge production function (dependent variable log 'arti')

	OLS	corrected OLS	WLS(und)	FGLS
K	0.094		0.094	0.102
	(0.022)***	(0.024)***	(0.022)***	(0.022)***
fac	0.897		0.896	0.841
	(0.056)***	(0.080)***	(0.057)***	(0.056)***
post	0.049		0.049	0.071
	(0.030)	(0.036)	(0.030)	(0.031)**
und	0.154		0.156	0.147
	(0.046)***	(0.056)***	(0.047)***	(0.040)***
tech	0.133		0.137	0.105
	(0.081)*	(0.072)*	(0.081)*	(0.050)**
edu	0.655		0.657	0.603
	(0.099)***	(0.105)***	(0.103)***	(0.098)***
time	0.058		0.053	0.033
	(0.030)*	(0.030)*	(0.029)*	(0.025)
cons	-2.832		-2.836	-2.714
	(0.304)***	(0.357)***	(0.308)***	(0.307)***
Adj R-sq	0.933		0.936	0.952
Breusch-Pagan	28.805***		24.705***	5.532

Significance levels: *** $p<0.01$; ** $p<0.05$; * $p<0.1$

5.4.2 Intramural R&D premium

Table 17 and Table 18 summarize the results of controlling for the importance of the different funding sources as modelled in equation (5.9). At first, many of coefficients for the share of funding sources are statistically insignificant except for the share of a university's intramural R&D expenditure when regressing on the total-journal articles. Intramural R&D investment seems a good solution to encourage scientific production. Although coefficients are not significant, there are several points to consider. Investments from private sector have a premium effect relative to government's funding. Intramural R&D also has premium effect. Foreign funding to university has premium effect for 'leading-journal' articles and discount effect for 'total-journal' articles. This result implies that university who wants to increase scientific production has to increase its intramural-R&D share.

Table 17 Knowledge production function including share of funding source (dependent variable log 'sci')

	OLS	corrected OLS	WLS(fac)	FGLS
K	0.609		0.620	0.665
	(0.054)***	(0.075)***	(0.052)***	(0.048)***
sPriv	-0.077		0.046	0.055
	(0.442)	(0.544)	(0.433)	(0.384)
sUniv	1.167		1.006	0.475
	(0.397)***	(0.546)**	(0.399)**	(0.366)
sFrgn	-0.067		-2.147	1.356
	(8.629)	(7.178)	(8.444)	(3.754)
fac	1.309		1.268	1.128
	(0.134)***	(0.165)***	(0.124)***	(0.120)***
post	-0.138		-0.124	-0.143
	(0.069)**	(0.092)	(0.068)*	(0.065)**
und	-0.148		-0.150	-0.054
	(0.108)	(0.116)	(0.101)	(0.099)
tech	0.683		0.631	0.593
	(0.186)***	(0.166)***	(0.184)***	(0.129)***
edu	0.219		0.133	0.245
	(0.242)	(0.348)	(0.263)	(0.356)
time	0.200		0.184	0.137
	(0.069)***	(0.069)***	(0.064)***	(0.052)***
cons	-11.574		-11.563	-12.109
	(0.712)***	(0.906)***	(0.698)***	(0.701)***
Adj R-sq	0.893		0.907	0.928
Breusch-Pagan	34.858***		25.8147***	8.7143

Significance levels: *** $p<0.01$; ** $p<0.05$; * $p<0.1$

Table 18 Knowledge production function including share of funding source (dependent variable log 'arti')

	OLS	corrected OLS	WLS(post)	FGLS
K	0.113		0.120	0.120
	(0.024)***	(0.024)***	(0.024)***	(0.022)***
sPriv	0.173		0.167	0.110
	(0.190)	(0.231)	(0.188)	(0.185)
sUniv	0.332		0.351	0.379
	(0.171)*	(0.171)*	(0.169)**	(0.168)**
sFrgn	-5.294		-5.207	-4.614
	(3.736)	(4.414)	(3.591)	(3.369)
fac	0.891		0.866	0.848
	(0.058)***	(0.080)***	(0.056)***	(0.053)***
post	0.044		0.049	0.071
	(0.030)	(0.035)	(0.031)	(0.031)**
und	0.135		0.142	0.120
	(0.047)***	(0.057)**	(0.045)***	(0.043)***
tech	0.104		0.103	0.084
	(0.081)	(0.071)	(0.079)	(0.075)
edu	0.600		0.604	0.576
	(0.104)***	(0.110)***	(0.102)***	(0.099)***
time	0.056		0.052	0.040
	(0.030)*	(0.030)*	(0.028)*	(0.025)
cons	-2.937		-3.013	-2.851
	(0.306)***	(0.364)***	(0.304)***	(0.300)***
Adj R-sq	0.934		0.941	0.954
Breusch-Pagan	28.494***		17.428*	4.890

Significance levels: *** $p<0.01$; ** $p<0.05$; * $p<0.1$

5.4.3 Knowledge production functions embedded with scientific collaboration

Table 19 and Table 20 are results for adding network position and core preference respectively to knowledge production function. Table 19 presents the results for equation (5.13), which estimates the model including network position variables. In terms of network position, closeness centrality has significantly positive effects on scientific production while betweenness centrality has significantly negative effects. However inserting this network position to our production function made some problem of interpreting result. It leaded to somewhat large reductions in elasticity of main variables, knowledge capital (from 0.59 to 0.42) and faculty size (from 1.21 to 0.96). This is caused by a new variable’s high correlation with existing variables, shown as Table 21. Therefore the network position is not valid for an independent variable in the knowledge production function.

On the other hand, core preference to USA seems to be much proper variable to control for external-university-world spillover effect. Its coefficient is positive but statistically insignificant, of 0.59, indicating that the more collaboration with US organizations, the more scientific production from Korean universities. Not only the core preference to USA is not a significant independent variable, but also it does not provide the information on external spillovers comparable with existing independent variables.

Therefore we concluded that both of centrality and core preference are not valid new variables for our knowledge production function.

Table 19 Knowledge production function including network centralities (dependent variable log 'sci')

	OLS	corrected OLS	WLS(post)	FGLS
K	0.313		0.352	0.417
	(0.054)***	(0.068)***	(0.053)***	(0.054)***
bet	-7.819		-6.906	-9.287
	(1.807)***	(2.113)***	(1.520)***	(2.347)***
clo	10.336		9.132	7.025
	(1.310)***	(1.609)***	(1.231)***	(1.138)***
fac	0.849		0.888	0.963
	(0.103)***	(0.116)***	(0.099)***	(0.095)***
post	-0.122		-0.138	-0.139
	(0.053)**	(0.073)*	(0.054)**	(0.056)**
tech	0.756		0.715	0.693
	(0.160)***	(0.156)***	(0.154)***	(0.114)***
time	0.227		0.211	0.188
	(0.067)***	(0.068)***	(0.064)***	(0.056)***
cons	-11.705		-11.688	-11.857
	(0.386)***	(0.432)***	(0.368)***	(0.394)***
Adj R-sq	0.898		0.910	0.915
Breusch-Pagan	28.381***		21.707***	13.247*

Significance levels: *** p<0.01; ** p<0.05; * p<0.1

Table 20 Knowledge production function including core preferences (dependent variable log 'sci')

	OLS	corrected OLS	WLS(fac)	FGLS
K	0.566		0.588	0.599
	(0.051)***	(0.071)***	(0.049)***	(0.046)***
Pus	0.839		0.569	0.587
	(0.463)*	(0.702)	(0.483)	(0.466)
fac	1.325		1.254	1.189
	(0.114)***	(0.135)***	(0.109)***	(0.094)***
post	-0.191		-0.166	-0.161
	(0.070)***	(0.101)*	(0.068)**	(0.063)**
tech	0.796		0.758	0.697
	(0.171)***	(0.122)***	(0.166)***	(0.107)***
edu	0.507		0.392	0.191
	(0.220)**	(0.333)	(0.240)	(0.299)
time	0.183		0.174	0.139
	(0.071)**	(0.072)**	(0.065)***	(0.052)***
cons	-11.888		-11.947	-11.758
	(0.457)***	(0.554)***	(0.428)***	(0.377)***
Adj R-sq	0.885		0.902	0.927
Breusch-Pagan	40.521***		28.430***	8.382

Significance levels: *** p<0.01; ** p<0.05; * p<0.1

Table 21 Correlation among variables

	articles	K	fac	post	bet	clo
K	0.941***					
fac	0.902***	0.802***				
post	0.911***	0.871***	0.903***			
bet	0.788***	0.851***	0.627***	0.672***		
clo	0.872***	0.816***	0.907***	0.870***	0.574***	
Pus	0.080	0.081	-0.001	0.088	0.078	0.003

Significance level: *** p<0.01

When we include the variable aiming to capture the knowledge spillovers (Table 22), the magnitude of the elasticity of the knowledge stock changes only marginally (remaining stable). The variable CU_{it} is highly significant and has a positive value for the estimated parameter of 0.35. That the value of the time dummy drops from 0.16 to 0.11 validates our conjecture that the time trend was capturing part of the spillover effects. This is same to the dummy for the technological universities; they used knowledge spillovers to some extent. Another meaningful result relates to the volume of the coefficients for knowledge capital and knowledge spillovers from collaborative research. Their sum (1.09) is above one suggesting the presence of increasing returns to scale at university-world level when we include knowledge spillover effects. Table 22 shows the result from estimating a model (5.16). The scientific output elasticities of knowledge capital, cross-university spillovers, and spillovers from the US are 0.59, 0.35, and 0.15, respectively. Comparing these values with ones from other studies (Crespi & Geuna, 2008), we found that spillover effect on the university world at Korea is a little small even when the two spillovers are summed up (0.49).

Table 22 Extended knowledge production function for Korean universities (dependent variable log 'sci')

	OLS	corrected OLS	WLS(K)	FGLS
K	0.545		0.551	0.594
	(0.052)***	(0.071)***	(0.051)***	(0.046)***
CU	0.330		0.317	0.349
	(0.121)***	(0.170)*	(0.118)***	(0.100)***
US	0.189		0.186	0.145
	(0.071)***	(0.107)*	(0.071)***	(0.072)**
fac	1.352		1.321	1.183
	(0.116)***	(0.134)***	(0.113)***	(0.094)***
post	-0.225		-0.206	-0.177
	(0.072)***	(0.103)**	(0.072)***	(0.063)***
tech	0.749		0.741	0.641
	(0.174)***	(0.122)***	(0.165)***	(0.104)***
edu	0.468		0.360	0.143
	(0.233)**	(0.361)	(0.237)	(0.320)
time	0.168		0.157	0.112
	(0.073)**	(0.075)**	(0.070)**	(0.052)**
cons	-21.715		-21.448	-21.158
	(2.428)***	(4.088)***	(2.403)***	(2.199)***
Adj R-sq	0.881		0.889	0.927
Breusch-Pagan	44.589***		39.100***	14.072*

Significance levels: *** p<0.01; ** p<0.05; * p<0.1

5.4.4 Comparison of knowledge spillovers

As listed in Table 23, elasticity in terms of knowledge spillovers is either larger or smaller than that of own R&D expenditure. Therefore it should be cautiously interpreted that spillovers are enough or not. In the meantime, the only work which uses scientific output as the main dependent variable is Crespi and Geuna (2008)'s investigation using international scientific co-authorship. At the country level, knowledge spillovers has a larger elasticity of scientific output than that of own R&D expenditures. Now we have a benchmark to compare.

Table 23 Elasticities of returns to own and others' R&D

Study	Sample	Period	Modelling of Spillovers	Elasticity of Own R&D	Elasticity of Spillover
Adams & Jaffe (1996)	US chemical sector : 21,546 plant-years	1974-88	Spatial correlation in R&D product fields	0.05	0.07
Los & Verspagen (2000)	US: 680 firms	1977-91	External R&D stocks ; weighted by patent flows	0.0-0.07	0.33-0.68
Ornaghi (2006)	Spain : more than 2000 manufacturing firms	1990-99	External R&D stocks ; weighted by size and industry proximity	0.10 (proc) 0.24 (prod)	0.02 (proc) 0.12 (prod)
Verspagen (1997)	14 countries, 22 sectors	1974-92	Patents, imports and indirect imports	0.10	0.03 (domestic) 0.05 (foreign)
Keller (1998)	22 countries	1971-90	Random weights interacted with import/GDP	0.13 (G7) 0.035 (others)	0.05 (G7 to others)
Kao et al. (1999)	22 countries	1971-90	Import share	0.21 (G7) 0.08 (others)	0.26

Van Pottelsberghe & Lichtenberg (2001)	13 countries	1971-90	Imports and outward or inward FDI	0.05 (import) 0.06 (outw. FDI) 0.08 (inw. FDI)	0.07 (import) 0.04 (outw. FDI) 0.01 (inw. FDI)
Luintel & Khan (2004)	10 countries	1965-99	Import share	0.29	0.12
Guellec & van Pottelsberghe (2004)	16 countries	1980-98	Technology-distance	0.13	0.46
Lopez-Pueyo et al. (2008)	6 countries, 10 manufacturing sectors	1979-2000	Same foreign industry and other foreign industry	0.14	0.11 (same) 0.21 (other)
Crespi & Geuna (2008)	14 countries	1981-2002	international scientific co-authorship	0.45 (publication) 0.50 (citation)	0.51 (publication) 0.60 (citation)

5.5 Conclusion

In this chapter, we developed a model of knowledge production function embedding a better understanding of the collaborative characteristics in scientific production process. The analysis was undertaken from a 'university world' viewpoint, based on a hypothesis that university should be affected by both an 'inside' and an 'outside' of the university world. We employed a sample of 141 Korean university schools for which we have information on R&D expenditure and scientific collaboration. As a proxy for output we used published scientific journal articles. Most inputs and outputs for this sample of universities are from the HEiK containing information recorded during 2008–2012. For the scientific collaboration, measured as co-authorship in the articles, the *1981–2010 South Korea NCR* (NCR) is utilized for collaboration characteristics of each university during 2006–2010. Given these data, we provided an assessment of the scientific returns to investment in university research, particularly considering the existence and characteristics of two spillovers: the one is 'cross-university' and the other is 'from US.' The results implicate that spillover effect has a real existence and a deep relation with scientific collaboration, while this effect is somewhat weak. The small spillover effect request either much more collaboration or much efficient co-operation.

The approach selected in this chapter is pragmatic employing quantitative methods to examine variables expected to influence scientific achievement. For specific, it develops econometric models founded in production function metaphor for relating a sub-set of inputs to university research output, publications in scientific journal. This chapter is not intent on searching for a "scientific constant," such as accurate estimates of the output elasticities with respect to investment. Instead, it aims to explore key features of the process of scientific production at Korean universities: (a) premium of intramural funding for university research and (b) the knowledge spillovers within university world and knowledge spillovers from outside the university world.

The model including the share of different funding source shows a significant premium effects of funding by university itself than one of funds from the government, private sectors, and foreign countries. The analysis of 'cross-university' spillovers indicates a significant impact on a universities' publications from the weighted R&D investments in other universities. Further, we investigated the spillovers 'from the US,' defined as a major core of the Korean scientific collaboration network in earlier chapter. The results were consistent with our conjectures: the one that spillovers among Korean universities are limited, and the other that spillovers from US are existent. For this model, economy of scale in Korean university science is attainable at the aggregate level (in contrast to the university-world level result).

From a strategy perspective our results highlight two important phenomena. First, for the higher scientific production, intramural R&D

expenditure has more effects than other funding sources. Autonomy given by intramural investments might promote a better performance in scientific production. Second, more collaboration with US organizations leads a university with more scientific publications. The scientific power of US has been confirmed by some works (Crespi & Geuna, 2008). Collectively, university has to increase intramural funded collaborative research with the US when it needs more scientific gains from R&D expenditure.

From a policy perspective the results also highlight two relevant circumstances. First, the elasticitiy of scientific output with respect to spillover effect among Korean universities is limited. This might be caused by the fragmented collaboration structure of 'university world' level. The fragmented collaboration appeared as an inter-tier tendency of Korean university in scientific partner selection (Ahn et al., 2014), and became visible as a star network after removing nodes of external world in Chapter 4. It is required that Korean universities recognize other universities in their world as to be partners rather than competitors, and that policy makers recognize this phenomenon as a problem we solve to make an efficient scientific system at the 'university level.' Second, that the impacts of US' science is a significant thing is considerable. There exists a hazard of Korean universities defining their science priorities without consideration for US science (Crespi & Geuna, 2008).

Chapter 6. Conclusion

6.1 Summary of Results

This study developed a model of knowledge production function that embedded a better understanding of the collaborative characteristics in the scientific production process. The analysis was undertaken from a “university world” perspective, based on the hypothesis that universities should be affected by institutions both “inside” and “outside” the university world. We employed a sample of Korean university schools, for which we have information on both their R&D expenditure and scientific collaboration.

Chapter 3 examined the characteristics of knowledge production, investigating scientific journal articles published by Korean universities in 1981–2010. Above all, scientific collaboration between universities is described as a cross-tier cooperative tendency, where schools prefer a different level of publication ranking rather than the same level for partners in collaborative research. The propensity ratios of tier combinations reveal that Korean scientists prefer inter-tier to intra-tier collaboration. This tendency could be the result of strategic partner selection by lower-tier universities, with the ultimate goal of achieving scientific success by generating higher-impact research.

Chapter 4 implies that the scientific collaboration network within the

Korean university sphere presents three characteristics: (a) A divided network that indicates less collaboration beyond national borders; (b) A core/peripheral network that reveals the huge importance of a few cores; (c) A star network that shows the fragmentary nature of collaboration among Korean universities.

In the fifth chapter, the model including shares of different funding sources shows the significant premium effects of funding by universities themselves rather than when receiving funding from the government, private sector, or foreign countries. The analysis of “cross university” spillovers indicates a significant impact on a university’s publications from the weighted R&D investments in other universities. Furthermore, we investigated the spillovers “from the US,” defined as a major core of the Korean scientific collaboration network in the previous chapter. The results were consistent with our conjecture: Spillovers among Korean universities are limited, and other spillovers, such as from the US exist. For this model, the economy of scale in Korean university science is attainable at the aggregate level (in contrast to the university-world level result).

The approach selected in Chapter 5 is the pragmatic employment of quantitative methods to examine the variables expected to influence scientific achievements. For specificity, it develops econometric models founded in production function metaphors to relate a subset of inputs to university research output which is the number of publications in scientific journals. This study is not intent on searching for “scientific constants” such as accurate estimates of output elasticities with respect to the level of investment. Instead, it aims to explore the key features of the process of scientific production in Korean universities: (a) Premium of intramural funding for university research and (b) The knowledge spillovers within the university world and knowledge spillovers from outside the university world.

6.2 Implications

Chapter 3 indicates that the Korean university sector is inefficient at producing high-impact publications. Collaborative research between high-tier organizations should be expanded.

From a strategy perspective, Chapter 4 highlights an important phenomenon. Universities are required to conduct more collaborative research with cores inside the network when they need to increase their scientific production. In the core/periphery structure, there is no special space; closeness to cores decides the power of a given actor. In the case of our study, the US seems a powerful option.

From a policy perspective, Chapter 4 also highlights two relevant circumstances. First, the structure of collaboration at the "university-world" level is a star-shaped network centered on the SNU. This structure is very inefficient for disseminating information to the whole network. The fragmented collaboration also appears as an inter-tier tendency of Korean universities in scientific partner selection (Ahn et al., 2014), and appears as a star network after the removal of external world nodes. It is required that Korean universities recognize other universities in their world as partners rather than competitors. Second, the USA is a core of the knowledge network for Korean universities. If we want a situation where one or more domestic universities are the core of the network, we again require more collaboration among universities.

Chapter 5 highlights two important phenomena. First, for higher scientific production, intramural R&D expenditure has a greater effect than other funding sources. Autonomy given by intramural investment might promote a better performance in scientific production. Second, greater collaboration with US organizations leads to universities with a greater number of scientific publications. The scientific power of the US has been confirmed by some works (Crespi & Geuna, 2008). Collectively, universities have to increase their intramural funds for collaborative research with the US when they require more scientific gains from R&D expenditure.

Chapter 5 also highlights two relevant circumstances. First, the elasticity of scientific output with respect to the spillover effect among Korean universities is limited. This might be caused by the fragmentary collaboration structure of the “university world” level. The fragmented collaboration appeared as an inter-tier tendency of Korean universities in scientific partner selection (Ahn et al., 2014), and has become visible as a star network after the removal of “external world” nodes in Chapter 4. It is required that Korean universities recognize other universities in their world as partners rather than competitors, and that policy makers recognize this phenomenon as a problem that should be solved to increase the efficiency of the scientific system at the "university level." Second, the impact of US science as a significant thing should be considered. There exists a hazard of Korean universities defining their science priorities without considering US science (Crespi & Geuna, 2008).

6.3 Contribution and Limitation

The first contribution of this study is applying the econometric method and perspective to describe the status of knowledge production by Korean universities (Chapter 3). Second, we used social network analysis to explore some of the properties of scientific collaboration networks, particularly their structural relationships (Chapter 4). The last contribution of this study is an extended model for relating scientific output to the universities' own knowledge capital, a cross-university spillover, and a spillover from the US (Chapter 5).

This study could not examine outputs other than knowledge (scientific results). The role of universities is to couple basic research and teaching. As countries progressively shift towards knowledge-based economies, there is a positive supply response on the part of universities to the increasing demand for basic knowledge and highly skilled people (Clark, 1993). Consideration of the students as an educational output, rather than as a control variable, may lead to other research questions.

Bibliography

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History*, 46, 385–406.

Adams, J. (1990). Fundamental stocks of knowledge and productivity growth. *The Journal of Political Economy*, 98, 673–702.

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: evidence from U.S. universities, 1981–1999. *Research Policy*, 34, 259–285.

Adams, J., & Griliches, Z. (1996). *Measuring science: an exploration* (NBER Working Paper 5478). Cambridge: National Bureau of Economic Research.

Adams, J., & Griliches, Z. (1998). Research productivity in a system of universities. *Annales d'Economie et de Statistique*, 49–50, 127–162.

Adams. J. D., & Jaffe, A. B. (1996). Bounding the effects of R&D: an investigation using matched establishment-frim data. *Rand Journal of Economics*, 27, 700–721.

Ahn, J., Oh, D., & Lee, J.-D. (2014). The scientific impact and partner selection in collaborative research at Korean universities. *Scientometrics*, 100, 173–188.

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: a longitudinal study. *Administrative Science Quarterly*, 45, 425–455.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286, 509–512.

Beaver, D. D. (2004). Does collaborative research have greater epistemic authority? *Scientometrics*, 60, 399–408.

Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative R&D and firm performance. *Research Policy*, 33, 1477–1492.

Ben-David, J. (1977). *Centers of learning: Britain, France, Germany, United States*. New York: McGraw-Hill Book Co.

Brass, D. J. (1984). Being in the right place: a structural analysis of individual influence in an organization. *Administrative Science Quarterly*, 29, 518–539.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47, 1287–1294.

Burt, R. S. (1987). Social contagion and innovation: cohesion versus structural equivalence. *The American Journal of Sociology*, 92, 1287–1335.

Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19, 1141–1164.

Choi, H. (2010). *Exploring the development of emerging sectors from the perspective of innovation systems: the case of the hydrogen energy and nano sectors in Korea* (Doctoral dissertation). Seoul: Graduate School of Seoul National Universtiy.

Clark, B. (1993). *The research foundations of graduate education: Germany, Britain, France, United States, Japan*. California: University of California Press.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48, 1–23.

Collins, H. M. (1974). The TEA set: tacit knowledge and scientific networks. *Science Studies*, 4, 165–186.

Cook, R. D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, 74, 169–174.

Crespi, G. (2007). The UK knowledge production function. In A. Bonaccorsi & C. Daraio (eds.), *Universities and strategic knowledge creation* (pp. 306–346). Cheltenham・Northampton: Edward Elgar.

Crespi, G., & Geuna, A. (2004). *The Productivity of Science* (Report prepared for the Office of Science and Technology, Department of Trade and Industry). Brighton: SPRU-University of Sussex.

Crespi, G. A., & Geuna, A. (2008). An empirical study of scientific production: a cross country analysis, 1981–2002. *Research Policy*, 37, 565–579.

Cronin, B., Shaw, D., & La Barre, K. (2003). A cast of thousands: coauthorship and subauthorship collaboration in the 20th century as manifested in the scholary journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, 54, 855–871.

DeBresson, C., & Amesse, F. (1991). Networks of innovators: a review and introduction to the issue. *Research Policy*, 20, 363–379.

Diestre, L., & Rajagopalan, N. (2012). Are all 'sharks' dangerous? New biotechnology ventures and partner selection in R&D alliances. *Strategic Management Journal*, 33, 1115–1134.

Eberhardt, M., Helmers, C., & Strauss, H. (2013). Do spillovers matter when estimating private returns to R&D? *The Review of Economics and Statistics*, 95, 436–448.

Everard, A., & Henry, R. (2002). A social network analysis of interlocked directorates in electronic commerce firms. *Electronic Commerce Research and Applications*, 1, 225–234.

Fagerberg, J., Fosaas, M., & Sapprasert, K. (2012a). Innovation: exploring the knowledge base. *Research Policy*, 41, 1132–1153.

Fagerberg, J., Landström, H., & Martin, B. R. (2012b). Exploring the emerging knowledge base of 'the knowledge society.' *Research Policy*, 41, 1121–1131.

Fairweather, J. S. (2005). Beyond the rhetoric: trends in the relative value of teaching and research in faculty salaries. *The Journal of Higher Education*, 76, 401–422.

Fleiss, J. L., & Berlin, J. A. (2009). Effect sizes for dichotomous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 237–253). New York: Russell Sage Foundation.

Foray, D. (2004). *The economics of knowledge*. Cambridge · London: MIT Press.

Foray, D., & Lissoni, F. (2010). University research and public-private interaction. In B. H. Hall & N. Rosenberg (eds.), *Handbook of the economics of innovation* (Vol. 1, pp. 275–314). Amsterdam · Oxford: North-Holland.

Fox, J. (2008). *Applied regression analysis and generalized linear models* (2nd ed.) Thousand Oaks: Sage Publications.

Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4, 540–553.

Freeman, C. (1991). Network of innovators: a synthesis of research issues. *Research Policy*, 20, 499–514.

Frenken, K., (2002). A new indicator of European integration and an application to collaboration in scientific research. *Economic Systems Research*, 14, 345–361.

Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: a case study of Harvard University's publications. *Scientometrics*, 87, 251–265.

Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*, 63, 323–335.

Greene, W. (2012). *Econometric analysis* (7th ed.). Upper Saddle River: Pearson Education.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10, 92–116.

Guellec, D., & van Pottelsberghe de la Potterie, B. (2001). *R&D and productivity growth: panel data analysis of 16 OECD countries* (OECD Science, Technology and Industry Working Papers 2001/03). Paris: OECD Publishing.

Guerrero-Bote, V. P., Olmeda-Gómez, C., & Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, 64, 392–404.

Gulati, R. (1998). Alliances and networks. *Strategic Management Journal*, 19, 293–317.

Hall, B. H., Mairesse, J., & Mohnen, P. (2010). Measuring the returns to R&D. In B. H. Hall & N. Rosenberg (eds.), *Handbook of the economics of innovation* (Vol. 2, pp. 1033–1082). Amsterdam・Oxford: North-Holland.

Hoch, P. K. (1987). Migration and the generation of new scientific ideas. *Minerva*, 25, 209–237.

Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010). Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39, 662–673.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics* (Vol. 1, pp. 221–233). Berkeley: University of California Press.

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review*, 79, 957–970.

Jenkins, A. (2004). *A guide to the research evidence on teaching-research relations*. York: The Higher Education Academy.

Johnston, J., & DiNardo, J. (1997). *Econometric methods* (4th ed.). New York: McGraw-Hill.

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: shifting impact, geography, and stratification in science. *Science*, 322, 1259–1262.

Kao, C., Chiang, M., & Chen, B. (1999). International R&D spillovers: an application of estimation and inference in panel cointegration. *Oxford Bulletin of Economics and Statistics*, 61, 691–709.

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31, 31–43.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.

Keller, W. (1998). Are international R&D spillovers trade related? Analyzing spillovers among randomly matched trade partners. *European Economic Review*, 42, 1469–1481.

Kim, K. (2008). *The structural evolution of the web2.0 service network under subgroup structures* (Doctoral dissertation). Seoul: Graduate School of Seoul National Universtiy.

Kim, K.-W. (2006). Measuring international research collaboration of peripheral countries: taking the context into consideration. *Scientometrics*, 66, 231–240.

Knoke, D., & Yang, S. (2008). *Social network analysis* (2nd ed.). Thousand Oaks: Sage Publications

Korfiatis, N. T., Poulos, M., & Bokos, G. (2006). Evaluating authoritative sources using social networks: an insight from Wikipedia. *Online Information Review*, 30, 252–262.

Landström, H., Harirchi, G., & Å ström, F. (2012). Entrepreneurship: exploring the knowledge base. *Research Policy*, 41, 1154–1181.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35, 673–702.

Lee, Y.-G. (2008). Patent licensability and life: a study of U.S. patents registered by South Korean public research institutes. *Scientometrics*, 75, 463–471.

Leydesdorff, L., & Sun, Y. (2009). National and international dimensions of the Triple Helix in Japan: university-industry-government versus international coauthorship relations. *Journal of the American Society for Information Science and Technology*, 60, 778–788.

Los, B., & Verspagen, B. (2000). R&D spillovers and productivity: evidence from U.S. manufacturing industries. *Empirical Economics*, 25, 127–148.

Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy*, 20, 1–12.

Marsh, H. W., & Hattie, J. (2002). The relation between research productivity and teaching effectiveness. *The Journal of Higher Education*, 73, 603–641.

Martin, B. R. (2012). Are universities and university research under threat? Towards an evolutionary model of university speciation. *Cambridge Journal of Economics*, 36, 543–565.

Martin, B. R., Nightingale, P., & Yegros-Yegros, A. (2012). Science and technology studies: exploring the knowledge base. *Research Policy*, 41, 1182–1204.

Monsuur, H. (2007). Stable and emergent network topologies: a structural approach. *European Journal of Operational Research*, 183, 432–441.

Narin, F., Hamilton, K., & Olivastro, D. (1997). The increasing linkage between US technology and public science. *Research Policy*, 26, 317–330.

Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papaers. *Scientometrics*, 21, 313–323.

Organisation for Economic Co-operation and Development. (2002). *Benchmarking industry-science relationships*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development. (2003). *The sources of economic growth in OECD countries*. Paris: OECD Publications.

Ornaghi, C. (2006). Spillovers in product and process innovation: evidence from manufacturing firms. *International Journal of Industrial Organization*, 24, 349–380.

Pavitt, K. (1991). What makes basic research economically useful? *Research Policy*, 20, 109–119.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., & Sobrero, M. (2013). Academic engagement and commercialisation: a review of the literature on university-industry relations. *Research Policy*, 42, 423–442.

Ramsden, P., & Moses, I. (1992). Associations between research and teaching in Australian higher education. *Higher Education*, 23, 273–295.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98, S71–S102.

Salter, A. J., & Martin, B. R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*, 30, 509–532.

Saxton, T. (1997). The effects of partner and relationship characteristics on alliance outcomes. *Academy of Management Journal*, 40, 443–461.

Scott, J. (1991). *Social network analysis: a handbook*. London: Sage Publication.

Smart, P., Bessant, J., & Gupta, A. (2007). Towards technological rules for designing innovation networks: a dynamic capabilities view. *International Journal of Operations & Production Management*, 27, 1069–1092.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39, 312–320.

Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology*, 41, 643–681.

Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics*, 81, 177–193.

Stephan, P. E. (2010). The economics of science. In B. H. Hall & N. Rosenberg (eds.), *Handbook of the economics of innovation* (Vol. 1, pp. 217–273). Amsterdam・Oxford: North-Holland.

Van der Valk, T., Meeus, M. T. H., Moors, E. H. M., & Faber, J. (2010). R&D collaboration in the life sciences: finding partners and the role of resource-based inducements and opportunities. *International Journal of Innovation Management*, 14, 179–199.

Van Pottelsberghe, B., & Lichtenberg, F. (2001). Does foreign direct investment transfer technology across borders? *Review of Economics and Statistics*, 83, 490–497.

Verspagen, B. (1997). Estimating international technology spillovers using technology flow matrices. *Review of World Economics*, 133, 226–248.

Waltman, L., Van Eck, N. J., Van Leeuven, T. N., Visser, M. S., Van Raan, A. F. J. (2011). Towards a new crown indicator: some theoretical considerations. *Journal of Informetrics*, 5, 37–47.

Wang, F., & Sun, Y. (2008). Self-organizing peer-to-peer social networks. *Computational Intelligence*, 24, 213–233.

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. New York: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393, 440–442.

Welsh, R., Glenna, L., Lacy, W., & Biscotti, D. (2008). Close enough but not too far: assessing the effects of university-industry research relationships and the rise of academic capitalism. *Research Policy*, 37, 1854–1864.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817–838.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data* (1st ed.). Cambridge・London: MIT Press.

Wooldridge, J. M. (2003). *Introductory econometrics: a modern approach* (2nd ed.). Cincinnati: South-Western College Publishing.

Wooldridge, J. M. (2008). *Introductory econometrics: a modern approach* (4th ed.). Cincinnati: South-Western College Publishing.

Wray, K. B. (2002). The epistemic significance of collaborative research. *Philosophy of Science*, 69, 150–168.

Appendix 1: Deriving the Basic Knowledge Production Function

Let $Y = F(X, K, u)$ be the production function which connects some measure of output to the inputs, where X stands for conventional inputs such as labor and capital. We define also $TFP = Y / X$ as the level of total factor productivity. Let us assume the production function to be Cobb-Douglas and rewrite F as:

$$Y = AL^{\alpha}C^{\beta}K^{\gamma}e^{u}$$

where A is a constant stands for technical progress.

Then we define a conventional total input index X as:

$$X = L^{s}C^{1-s}$$

where s is the observed factor share of labor input. When we assume that s is proportional to the true coefficient of labor, we shall obtain $s = \alpha / (\alpha + \beta)$. We then also get measured total factor productivity as:

$$TFP = Y / X = AX^{\alpha+\beta-1}K^{\gamma}e^{u}$$

The constant returns to scale (CRS) assumption leads to $TFP = G(K, \mathrm{u})$ as:

$$TFP = AK^{\gamma}e^{u}$$

As scientific output is also “intermediate (Adams & Griliches, 1998)” such as productivity, we might use same production function (from Adams & Griliches, 1996; to Crespi & Geuna, 2008).

Abstract (Korean)

본 연구는 지식 생산의 과정을 탐구하는 데 그 목적을 가진다. 이를 위해 과학계량학적 관점, 계량경제학적 방법론, 사회 연결망의 견해, 성장 이론의 생삼함수 접근법 등을 차용하였다. 이들을 종합적으로 이용하여 연구협력의 특성을 포함한 지식 생산 과정의 성격을 규명하는 데 치중하였다.

첫째로, 최상위 대학을 포함한 대학 간 협력이 영향력 높은 학술지 논문을 생산함을 설명하였다. 또한 우수 대학들은 하위 대학들과의 협력 시 영향력에 있어 손해를 보는 반면, 하위 파트너들은 협력을 통한 영향력 이득을 취함을 밝혔다. 한국 대학들은 같은 층위의 대학들과 협력하지 않는 경향성을 보였다.

둘째로, 공저자 행위와 같은 상호 관계를 연구하는 데 적합한 사회연결망분석(SNA)을 사용하여 한국 대학들로 구성된 연구협력망 내에서의 구조 관계를 그려냈다. 이러한 목적으로 국내 비 대학 기관들과 외국 기관들을 묶어 ‘대학 외부’라 정의하고, 이를 ‘대학 내부’와 구분지어 비교하였다. 그런 다음 한국 학계의 학술지 공저자 연결망 내에서의 구조 관계를 조사하기 위한 연결망 그림을 완성하였다.

셋째로, 학술지 생산 과정에서의 협력적 특성을 고려한 지식생산함수의 모형을 개발하였다. ‘대학’의 입장에서 바라보면 연구 활동이 대학의 ‘내부’뿐 아니라 ‘외부’에서도 영향을 받는다는

가정 하에 지식의 생산 과정을 모형화 하였다. 세부적으로, 학술지 단위의 연구개발 투자수익을 결과로 제공하는 지식생산함수를 ‘대학 간’ 스필오버 및 ‘미국 발’ 스필오버의 두 가지 스필오버 효과를 포함하여 확장하였다. 실증분석의 결과, 한국 대학의 지식 생산 과정에서 스필오버 효과는 실존하며 연구협력과 관계가 깊은 것으로 드러났다. 또한 연구개발의 투자수익에 비해 상대적으로 스필오버 효과가 작게 나타난 것은 국내 대학들의 협력이 양적·질적으로 부족한 현실을 반영한 것이라 볼 수 있기에 연구협력을 유도하는 다양한 정책이 필요한 실정이라 할 수 있겠다.

주요어 : 지식생산함수, 연구협력, 대학 연구활동, 파트너 선정,
공저자 연결망, 지식 스필오버

학 번 : 2005 - 21293